



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 116827

TO: Fozia Hamud
Location: rem/4d64/4c70
Art Unit: 1647
Monday, March 15, 2004

Case Serial Number: 09/997573

From: Edward Hart
Location: Biotech-Chem Library
REM-1A55
Phone: 571-272-2512

edward.hart@uspto.gov

Search Notes

Examiner Hamud,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

STIC-Biotech/ChemLib

116829

From: Hamud, Fozia
Sent: Friday, March 12, 2004 4:30 PM
To: STIC-Biotech/ChemLib
Subject: sequence search for 09/997,573

Kindly search SEQ IID NO:418 of 09/997,573 against commercial and interference data abses. Thanks.

Fozia Hamud
Patent Examiner
Art Unit 1647
Remsen, Room 4D64
Mail Box Remsen-4C70
272-0884

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 3/15/04
Date Completed: 3/15/04
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: 1
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: ASP
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 08:19:27 ; Search time 60 Seconds
(without alignments)
932.407 Million cell updates/sec

Title: US-09-997-573-418
Perfect score: 1031
Sequence: 1 MATLWGGLRLGSLLSLCL.....WKLQVQRKSVFDRHVLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001s:*
5: geneseq2002s:*
6: geneseq2003as:*
7: geneseq2003bs:*
8: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1031	100.0	198	3 AAY88275	Aay88275 Human TAN
2	1031	100.0	198	3 AAY66762	Aay66762 Membrane-
3	1031	100.0	198	3 AAY87231	Aay87231 Human sig
4	1031	100.0	198	3 AAY78807	Aay78807 Hydrophob
5	1031	100.0	198	4 AAM93740	Aam93740 Human pol
6	1031	100.0	198	4 AAB50966	Aab50966 Human PRO
7	1031	100.0	198	4 AAB20120	Aab20120 Human imm
8	1031	100.0	198	4 AAM38735	Aam38735 Human pol
9	1031	100.0	198	4 AAB50926	Aab50926 Human PRO
10	1031	100.0	198	4 AAB565285	Aab565285 Human PRO
11	1031	100.0	198	5 ABP61428	Abp61428 Human NF-
12	1031	100.0	198	6 ABUS58100	Abu58100 Human PRO
13	1031	100.0	198	6 ABUS59178	Abu59178 Novel hum
14	1031	100.0	198	6 ABUS2890	Abu2890 Human sec
15	1031	100.0	198	6 ABUS60609	Abu60609 Human sec
16	1031	100.0	198	6 ABUS13991	Abu13991 Human PRO
17	1031	100.0	198	6 ABUS72576	Abu72576 Novel hum
18	1031	100.0	198	6 ABUS71432	Abu71432 Human neo
19	1031	100.0	198	6 ABUS59325	Abu59325 Human sec
20	1031	100.0	198	6 ABUS26022	Abu26022 Human PRO
21	1031	100.0	198	6 ABUS59031	Abu59031 Human sec
22	1031	100.0	198	6 ABUS2409	Abu2409 Novel hum
23	1031	100.0	198	6 ABUS59474	Abu59474 Novel hum
24	1031	100.0	198	6 ABUS2240	Abu2240 Novel hum
25	1031	100.0	198	6 ABUS10946	Abu10946 Human PRO

26	1031	100.0	198	6 ABUS1698	Abu1698 Novel hum
27	1031	100.0	198	6 ABUS8637	Abu8637 Human sec
28	1031	100.0	198	6 ABO34151	Abu34151 Human PRO
29	1031	100.0	198	6 ADA37929	Ada37929 Human sec
30	1031	100.0	198	6 ADA21615	Ada21615 Human sec
31	1031	100.0	198	6 ADA10402	Ada10402 Human sec
32	1031	100.0	198	6 ADA17946	Ada17946 Human PRO
33	1031	100.0	198	6 ADA28054	Ada28054 Human sec
34	1031	100.0	198	6 ADA94634	Ada94634 Human sec
35	1031	100.0	198	6 ADA38859	Ada38859 Human sec
36	1031	100.0	198	6 ADA92980	Ada92980 Human sec
37	1031	100.0	198	7 ABO53237	Abu53237 Human sec
38	1031	100.0	198	7 ADA22541	Ada22541 Human sec
39	1031	100.0	198	7 ABO22607	Abu22607 Human sec
40	1031	100.0	198	7 ADA06707	Ada06707 Human sec
41	1031	100.0	198	7 ADA39400	Ada39400 Human sec
42	1031	100.0	198	7 ADB96426	Adb96426 Human PRO
43	1031	100.0	198	7 ADCS7898	Adcs7898 Human PRO
44	1031	100.0	198	7 ADCS5262	Adcs5262 Human PRO
45	1031	100.0	198	7 ADC12129	Adc12129 Human sec

ALIGNMENTS

RESULT 1

AAY88275

ID AAY88275 standard; protein; 198 AA.

AC AAY88275;

DT 16-OCT-2000 (first entry)

DE Human TANGO 184 protein.

XX TANGO 180; TANGO 181; TANGO 182; TANGO 183; TANGO 184; TANGO 185;

XX TANGO 186; TANGO 188; TANGO 189; TANGO 215; TANGO 187; human; murine;

XX secreted protein; transmembrane protein; gene therapy; vaccine;

XX diagnosis; treatment; detection.

XX Homo sapiens.

XX WO200018904-A2.

XX 06-APR-2000.

XX 30-SEP-1999; 99WO-US022817.

XX 30-SEP-1998; 98US-00164220.

XX 02-OCT-1998; 98US-00164169.

XX (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.

XX Barnes TM;

XX WPI; 2000-293144/25.

XX N-PSDB; AAA39941, AAA39942.

XX Isolated nucleic acids encoding TANGO polypeptides useful for preventing,

XX diagnosing and treating diseases associated with inappropriate protein

XX expression.

XX Claim 9; Fig 9; 249pp; English.

XX This invention describes novel human and murine nucleic acids encoding

XX TANGO polypeptides (which are either wholly secreted or transmembrane

XX proteins) which can be used for gene therapy and/or vaccination. The

XX peptides are designated TANGO 180 to TANGO 189 and TANGO 215. The nucleic

XX acids may be used to produce TANGO 180 to TANGO 189 and TANGO 215

XX polypeptides according to standard recombinant DNA methodologies. They

XX may also be used to detect and quantify the presence of TANGO nucleic

XX acids in a sample and therefore identify or diagnose diseases associated

XX with inappropriate TANGO expression (e.g. diseases related to over or

Sequence
'A'

See over

Sequence, Comparison

under expression of the polypeptides or the expression of inactive polypeptides). The nucleic acids and the polypeptides they encode may be used according to standard gene therapy protocols, to treat diseases associated with inappropriate TANGO expression by supplementing a patient's own production of the polypeptide of to rectify mutations that may result in expression of an abnormally active polypeptide. The polypeptides may also be used to identify and produce agonists and antagonists of TANGO expression and activity which may be used to modulate TANGO related processes and diseases. The polypeptides are particularly useful for use as antigens for producing antibodies to TANGO proteins which may be used for inhibiting the activity of TANGO proteins. They may also be used to detect and quantify the presence of TANGO proteins in samples and therefore identify patients in whom the protein is over- or under-expressed. This sequence represents the human TANGO 184 protein described in the method of the invention

```

SQ      Sequence 198 AA;

Query Match      100.0%; Score 1031; DB 3; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1 MATLWGGLLRLGSLLSLCSLALSLLLAQLSDAAKNFEDYRCKICPPYKENSGHYNNK 60
Db      1 MATLWGGLLRLGSLLSLCSLALSLLLAQLSDAAKNFEDYRCKICPPYKENSGHYNNK 60

QY      61 ISQKDCDCLHVVPEMPVRGPDVEAYCLRCECKYERSSVTIKVITIIYLSILGLLLYMW 120
Db      61 ISQKDCDCLHVVPEMPVRGPDVEAYCLRCECKYERSSVTIKVITIIYLSILGLLLYMW 120

QY      121 YLTLYVEPILKRLFGHAQLISDDDDIGDHQPFANAHDVLARSRRANVLNKVEYAQRWK 180
Db      121 YLTLYVEPILKRLFGHAQLISDDDDIGDHQPFANAHDVLARSRRANVLNKVEYAQRWK 180

QY      181 LQVQEQRKSVFDRHVVL 198
Db      181 LQVQEQRKSVFDRHVVL 198

```

	RESULT 2		
XX	AAY66762	IID AAY66762 standard; protein; 198 AA.	
XX	TAC	AAY66762;	
XX	OS	05-APR-2000 (first entry)	
DE	XW	Membrane-bound protein PROI375.	
XX	KX	Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;	
KW	KK	pharmaceutical; receptor immunoadhesin; gene mapping.	
XX	OS	Homo sapiens.	
XX	PV	WO9963088-A2.	
PD	PN	09-DEC-1999.	
XX	PF	02-JUN-1999;	
XX	PR	98US-USO12252.	
PR	PP	02-JUN-1998;	
PR	PR	98US-0087607P.	
PR	PR	98US-0087609P.	
PR	PR	98US-0087759P.	
PR	PR	98US-0087827P.	
PR	PR	98US-0088021P.	
PR	PR	98US-0088025P.	
PR	PR	98US-0088028P.	
PR	PR	98US-0088029P.	
PR	PR	98US-0088030F.	
PR	PR	98US-0088033P.	
PR	PR	98US-0088326P.	
PR	PR	98US-0088167P.	
PR	PR	98US-0088202P.	

Sequence Comparison

07-JUL-1998; 98US-0091978P.
07-JUL-1998; 98US-0091982P.
09-JUL-1998; 98US-0092182P.
10-JUL-1998; 98US-0092472P.
20-JUL-1998; 98US-0093339P.
30-JUL-1998; 98US-0094651P.
04-AUG-1998; 98US-0095282P.
04-AUG-1998; 98US-0095285P.
04-AUG-1998; 98US-0095301P.
04-AUG-1998; 98US-0095302P.
04-AUG-1998; 98US-0095318P.
04-AUG-1998; 98US-0095321P.
04-AUG-1998; 98US-0095325P.
10-AUG-1998; 98US-0095916P.
10-AUG-1998; 98US-0095929P.
11-AUG-1998; 98US-0096012P.
11-AUG-1998; 98US-0096143P.
11-AUG-1998; 98US-0096146P.
17-AUG-1998; 98US-0096329P.
17-AUG-1998; 98US-0096757P.
17-AUG-1998; 98US-0096766P.
17-AUG-1998; 98US-0096768P.
17-AUG-1998; 98US-0096773P.
17-AUG-1998; 98US-0096791P.
17-AUG-1998; 98US-0096867P.
17-AUG-1998; 98US-0096891P.
17-AUG-1998; 98US-0096894P.
17-AUG-1998; 98US-0096895P.
17-AUG-1998; 98US-0096897P.
18-AUG-1998; 98US-0096949P.
18-AUG-1998; 98US-0096950P.
18-AUG-1998; 98US-0096959P.
18-AUG-1998; 98US-0096960P.
18-AUG-1998; 98US-0097022P.
19-AUG-1998; 98US-0097141P.
20-AUG-1998; 98US-0097218P.
24-AUG-1998; 98US-0097661P.
26-AUG-1998; 98US-0097951P.
26-AUG-1998; 98US-0097952P.
26-AUG-1998; 98US-0097954P.
26-AUG-1998; 98US-0097955P.
26-AUG-1998; 98US-0097971P.
26-AUG-1998; 98US-0097974P.
26-AUG-1998; 98US-0097978P.
26-AUG-1998; 98US-0097979P.
26-AUG-1998; 98US-0097986P.
26-AUG-1998; 98US-0098014P.
31-AUG-1998; 98US-0098525P.
16-SEP-1998; 98US-0100634P.
12-JAN-1999; 99US-0115565P.

(GETH) GENENTECH INC.

Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
Wood WI, Yuan J;

WPI; 2000-072883/06.
N-PSDB; AA265108.

Membrane-bound proteins and related nucleotide sequences.

Claim 12; Fig 300; 822pp; English.

The invention provides membrane-bound PRO polypeptides and polynucleotides encoding them. The PRO sequences of the invention were identified based on extracellular domain homology screening. The PRO sequences have homology with proteins including LDL receptors, TIE ligands and various enzymes. The membrane-bound proteins and receptor molecules are useful as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be used as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. The PRO encoding sequences

CC are useful as hybridization probes, in chromosome and gene mapping and in the generation of antisense RNA and DNA. PRO nucleic acid sequences will also be useful for the preparation of PRO polypeptides, especially by recombinant techniques

XX Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 3; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGLLRGLSLISCLALSLLAQLSDAAKNFEDVRCCKICPPYKENSCHIYNKN 60
DB 1 MATLWGLLRGLSLISCLALSLLAQLSDAAKNFEDVRCCKICPPYKENSCHIYNKN 60
QY 61 ISQKDCDCLHVVEPMPVGRPDVEAYCLRCCKEYERSVTKVIIIVLSILGLLLYMV 120
DB 61 ISQKDCDCLHVVEPMPVGRPDVEAYCLRCCKEYERSVTKVIIIVLSILGLLLYMV 120
QY 121 YLTIVPEILKRLFGHAQLIQSDDDIGDHQPFANAHVLAHSRANVKNVEYAQQRWK 180
DB 121 YLTIVPEILKRLFGHAQLIQSDDDIGDHQPFANAHVLAHSRANVKNVEYAQQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
DB 181 LQVQQRKSVFDRHVLS 198

RESULT 3
AAV87231

ID AAV87231 standard; protein; 198 AA.

XX AAV87231;

XX DT 11-MAY-2000 (first entry)

XX Human signal peptide containing protein HSPB-8 SEQ ID NO:8.

Human; signal peptide-containing protein; HSPB; diagnosis; cancer; inflammation; cardiovascular disease; anticancer; anti-inflammatory; antimicrobial; nontoxic; neuroprotective; cardiovascular; hepatotropic; antasthmatic; gene therapy; cell proliferation; neurological disorder; KW reproductive disorder; developmental disorder; arteriosclerosis; KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia; asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia; Parkinson's disease; Huntington's diseases; ovulatory defect; muscular dystrophy.

XX Homo sapiens.

XX WC2000000610-A2.

XX 06-JAN-2000.

XX 25-JUN-1999; 99WO-US014484.

XX 26-JUN-1998; 98US-0090762P.

XX 31-JUL-1998; 98US-0094983P.

XX 01-OCT-1998; 98US-0102686P.

XX 11-DEC-1998; 98US-0112129P.

XX (INCY-) INCYTE PHARM INC.

XX Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;
XX Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;
XX Bandman O;

XX WPI; 2000-160673/14.

XX N-PSDB; AA298116.

XX New human signal peptide-containing proteins useful in treatment, prevention and diagnosis of e.g. cancer, inflammation and cardiovascular disease.

Sequence Comparison

See over

Sequence Comparison

XX AC AAM93740;
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide, SEQ ID NO: 3711.
XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX OS Homo sapiens.
XX PN EPI130094-A2.
XX PD 05-SEP-2001.
XX PF 07-JUL-2000; 2000EP-00114089.
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PR 02-MAY-2000; 2000JP-00183765.
XX PA (HELI-) HELIX RES INST.
XX OTa T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX PR WPI; 2001-524255/58.
XX DR N-PSDB; AAK94692.
XX XX 830 Primers useful for synthesizing full length cDNA clones and their use
XX PT in genetic manipulation.
XX PS Claim 8; SEQ ID NO 3711; 1380pp + Sequence Listing; English.
XX CC The invention relates to primers for synthesizing full length cDNA
XX CC clones. 830 cDNA molecules encoding a human protein have been isolated
XX CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX CC been determined. Primers for synthesizing the full length cDNA are useful
XX CC for clarifying the function of the protein encoded by the cDNA. The full
XX CC length clones were obtained by construction of full length enriched cDNA
XX CC libraries that were synthesised by the oligo-capping method. The primers
XX CC enable the production of the full length cDNA easily without any special
XX CC methods. The present sequence is a polypeptide encoded by a full length
XX CC human cDNA of the invention. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in CD-ROM
XX CC format directly from EPO
XX SQ Sequence 198 AA;
Query Match 100.0%; Score 1031; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATLWGGLRLGSLLSCLALS SVLLLAQLSDAAKNFEDVRCKICPPYKENS GHYNNK 60
DB 1 MATLWGGLRLGSLLSCLALS SVLLLAQLSDAAKNFEDVRCKICPPYKENS GHYNNK 60
QY 61 ISQKDCDCHLVVEPVPDPVPCVAYCLRCCKEYERSSVTIKVTIIIVLSIIGLILLYW 120
DB 61 ISQKDCDCHLVVEPVPDPVPCVAYCLRCCKEYERSSVTIKVTIIIVLSIIGLILLYW 120
QY 121 YLTLPVPIKRLFGHAQLIQSDDDIGDHPANAHDLVLRSRANVINKVEYAQRWK 180
DB 121 YLTLPVPIKRLFGHAQLIQSDDDIGDHPANAHDLVLRSRANVINKVEYAQRWK 180
QY 181 LQVQEQKSVFDRHVLS 198
DB 181 LQVQEQKSVFDRHVLS 198

RESULT 6
AAB50966
ID AAB50966 standard; protein; 198 AA.

XX AC AAB50966;
XX DT 21-MAR-2001 (first entry)
XX DE Human PRO1375 protein.
XX KW Human; PRO; cytostatic; nootropic; neuroprotective; respiratory general;
XX KW antiinflammatory; antiangiogenic; immunosuppressive; immunostimulant;
XX KW PRO agonist; cancer; inflammatory disorder; immunological disorder.
XX OS Homo sapiens.
XX PN WO200073348-A2.
XX PD 07-DEC-2000.
XX PF 30-MAY-2000; 2000WO-US014941.
XX PR 02-JUN-1999; 99WO-US012352.
XX PR 22-JUN-1999; 99US-0140650P.
XX PR 23-JUN-1999; 99US-0141037P.
XX PR 20-JUL-1999; 99US-0144758P.
XX PR 01-SEP-1999; 99WO-US020111.
XX PR 08-SEP-1999; 99WO-US020594.
XX PR 29-OCT-1999; 99US-0182506P.
XX PR 30-NOV-1999; 99WO-US028313.
XX PR 01-DEC-1999; 99WO-US028634.
XX PR 02-DEC-1999; 99WO-US028551.
XX PR 09-DEC-1999; 99US-0170262P.
XX PR 16-DEC-1999; 99WO-US030095.
XX PR 20-DEC-1999; 99WO-US030999.
XX PR 06-JAN-2000; 2000WO-US000376.
XX PR 11-FEB-2000; 2000WO-US003565.
XX PR 18-FEB-2000; 2000WO-US004341.
XX PR 19-FEB-2000; 2000WO-US004342.
XX PR 02-MAR-2000; 2000WO-US005841.
XX PR 03-MAR-2000; 2000US-0187202P.
XX PR 10-MAR-2000; 2000WO-US006319.
XX PR 15-MAR-2000; 2000WO-US006884.
XX PR 30-MAR-2000; 2000WO-US008439.
XX PR 17-MAY-2000; 2000WO-US013705.
XX PA (GETH) GENENTECH INC.
XX BK Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W, Kabakoff RC;
XX PI Shelton DL, Smith V, Watanabe CK, Wood WI;
XX XX WPI; 2001-016509/02.
XX DR N-PSDB; AAC91568.
XX XX Twenty eight nucleic acids encoding PRO polypeptides which are useful for
XX PT treating various tumors, e.g. breast cancer, and other inflammatory,
XX PT angiogenic and immunological disorders.
XX PS Claim 31; Fig 32; 189pp; English.
XX XX The present sequence is one of twenty eight novel PRO polypeptides. The
XX CC PRO polypeptides and their agonists, including antibodies, peptides, and
XX CC small molecule agonists, may be used to treat various tumors, e.g.,
XX CC cancers such as breast cancer, ovarian cancer, renal cancer, colorectal
XX CC cancer, uterine cancer, prostate cancer, lung cancer, bladder cancer,
XX CC central nervous system cancer, melanoma or leukaemia. They are also
XX CC useful for treating other disorders such as neuronal, glial, astrocytal,
XX CC hypothalamic and other glandular, macrophagal, epithelial, stromal and
XX CC blastocoeleic disorders, and inflammatory, angiogenic and immunological
XX CC disorders
XX XX Sequence 198 AA;
SQ Query Match 100.0%; Score 1031; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGLLRLGSLLSCLALSLLAQLSDAAKFNEDVRCCKICPPYKENSCHIYKN 60
Db 1 MATLWGLLRLGSLLSCLALSLLAQLSDAAKFNEDVRCCKICPPYKENSCHIYKN 60
QY 61 ISQKDCDCLHVPEPMPVGRGPDVEAYCLRCCKEYERSVTKVTIIYLSILGLLLYV 120
Db 61 ISQKDCDCLHVPEPMPVGRGPDVEAYCLRCCKEYERSVTKVTIIYLSILGLLLYV 120
QY 121 YLTLPVPEILKRLFGHAQLIQSDDDIGDHQPFANAHDLARSRLANVKNVEYAQRWK 180
Db 121 YLTLPVPEILKRLFGHAQLIQSDDDIGDHQPFANAHDLARSRLANVKNVEYAQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
Db 181 LQVQQRKSVFDRHVLS 198
RESULT 7
ID AAB20120 standard; protein; 198 AA.
XX
AC AAB20120;
XX
DT 30-APR-2001 (first entry)
XX
DE Human immunostimulant PRO1375.
XX
KW PRO1375; UNQ712; human; immune disease; autoimmune disease;
KW antirheumatic; antiarthritic; antiinflammatory; antianemic;
KW immunosuppressive; antithyroid; antidiabetic; neuroprotective;
KW hepatotropic; virucide; dermatological; antipsoriatic; antiasthmatic;
KW antiallergic; immunostimulant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 11..28
FT Modified-site /note= "transmembrane domain type I"
FT Modified-site /note= "N-myristoylation site"
FT Modified-site 60..64
FT Modified-site /note= "Asn is N-glycosylated"
FT Modified-site 78..86
FT Modified-site /note= "tyrosine kinase phosphorylation site"
FT Domain 103..125
FT /note= "transmembrane domain"
XX
PN WO200105972-A1.
XX
PD 25-JAN-2001.
XX
PF 15-MAR-2000; 2000WO-US006884.
XX
PR 20-JUL-1999; 99US-0144758P.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Fong S, Goddard A, Godowski PJ;
PI Gurney AL, Hillan KJ, Mark MR, Marsters SA, Pitti RM, Tumas D;
PI Watanabe CK, Wood WI;
XX
DR WPI; 2001-103149/11.
DR N-PSDB; AAF30062.
XX
PT New PRO polypeptides, nucleic acids and (ant)agonists, useful for
PT diagnosing and treating immune-related disorders, such as multiple
PT sclerosis, rheumatoid arthritis and diabetes.
XX
PS Claim 20; Fig 26; 127pp; English.
XX
CC The present sequence is that of novel human immunomodulator PRO1375
CC (UNQ712), as deduced from cDNA (see AAF30062) isolated from a human

CC pancreas tissue library. PRO1375 has a mol.wt. of 22 kDa and a pI of
CC 8.47. The invention provides polynucleotides (see AAF30050-62) encoding
CC novel human PRO proteins (see AAB20108-20) including PRO1375. Claimed
CC compositions comprising these proteins or their agonists are useful for
CC increasing infiltration of inflammatory cells into a tissue of a mammal,
CC stimulating or enhancing an immune response, or increasing the
CC proliferation of T-lymphocytes in a mammal in response to an antigen.
CC Claimed compositions comprising a PRO polypeptide or its antagonist have
CC the opposite effect. A claimed method for treating an immune related
CC disorder, such as a T cell disorder, involves administering a PRO
CC polypeptide, an agonist antibody or an antagonist antibody. The disorder
CC is selected from systemic lupus erythematosus, rheumatoid arthritis,
CC osteoarthritis, juvenile chronic arthritis, spondyloarthritis, systemic
CC sclerosis, idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, demyelinated diseases (such as multiple sclerosis), autoimmune
CC chronic active hepatitis, primary biliary cirrhosis, granulomatous
CC hepatitis, sclerosing cholangitis, inflammatory bowel disease (ulcerative
CC colitis and Crohn's disease), gluten-sensitive enteropathy, Whipple's
CC disease, (auto)immune-mediated skin diseases (such as bullous skin
CC disease, erythema multiforme and psoriasis), allergic diseases (such as
CC asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity and
CC urticaria), immunologic diseases of the lung and transplantation
CC associated diseases (such as graft rejection and graft-versus-host
CC disease) (all claimed). Claimed methods of diagnosing these disorders
CC comprise detecting the level of expression of the PRO gene. Also claimed
CC are a method of identifying a compound capable of inhibiting the
CC expression or activity of the PRO polypeptide, vectors, host cells,
CC antibodies, and a method of stimulating the proliferation of T
CC lymphocytes using PRO1375
SQ Sequence 198 AA;
Query Match 100.0%; Score 1031; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATLWGLLRLGSLLSCLALSLLAQLSDAAKFNEDVRCCKICPPYKENSCHIYKN 60
Db 1 MATLWGLLRLGSLLSCLALSLLAQLSDAAKFNEDVRCCKICPPYKENSCHIYKN 60
QY 61 ISQKDCDCLHVPEPMPVGRGPDVEAYCLRCCKEYERSVTKVTIIYLSILGLLLYV 120
Db 61 ISQKDCDCLHVPEPMPVGRGPDVEAYCLRCCKEYERSVTKVTIIYLSILGLLLYV 120
QY 121 YLTLPVPEILKRLFGHAQLIQSDDDIGDHQPFANAHDLARSRLANVKNVEYAQRWK 180
Db 121 YLTLPVPEILKRLFGHAQLIQSDDDIGDHQPFANAHDLARSRLANVKNVEYAQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
Db 181 LQVQQRKSVFDRHVLS 198
RESULT 8
AAB38735
ID AAB38735 standard; protein; 198 AA.
XX
AC AAB38735;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 1880.
XX
KW Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 08:19:27 ; Search time 60 Seconds
(without alignments)
932.407 Million cell updates/sec

Title: US-09-997-573-418
Perfect score: 1031
Sequence: 1 MARLMGGLRLGSLLSLCL.....WKLQVQEKSKSVFDRHVLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1031	100.0	198	3	AY88275 Human TAN
2	1031	100.0	198	3	AY66762 Membrane-
3	1031	100.0	198	3	AY87231 Human sig
4	1031	100.0	198	3	AY78807 Hydrophob
5	1031	100.0	198	4	AM93740 Human pol
6	1031	100.0	198	4	AB50966 Human PRO
7	1031	100.0	198	4	AB20120 Human imm
8	1031	100.0	198	4	AM38735 Human pol
9	1031	100.0	198	4	AB50926 Human PRO
10	1031	100.0	198	4	AB65285 Human PRO
11	1031	100.0	198	5	ABP61428 Human NF-
12	1031	100.0	198	5	ABU58100 Human PRO
13	1031	100.0	198	6	ABU59178 Novel hum
14	1031	100.0	198	6	ABU82690 Human sec
15	1031	100.0	198	6	ABU60609 Human sec
16	1031	100.0	198	6	ABU13991 Human PRO
17	1031	100.0	198	6	ABU72576 Novel hum
18	1031	100.0	198	6	ABU71432 Human neo
19	1031	100.0	198	6	ABU59325 Human sec
20	1031	100.0	198	6	ABO26022 Human PRO
21	1031	100.0	198	6	ABU59031 Human sec
22	1031	100.0	198	6	ABU92409 Novel hum
23	1031	100.0	198	6	ABU5474 Novel hum
24	1031	100.0	198	6	ABU92240 Novel hum
25	1031	100.0	198	6	ABU10946 Human PRO

26	1031	100.0	198	6	ABU81698	Abu81698 Novel hum
27	1031	100.0	198	6	ABU88637	Abu88637 Human sec
28	1031	100.0	198	6	ABO34151	Abu34151 Human PRO
29	1031	100.0	198	6	ADA37929	Ada37929 Human sec
30	1031	100.0	198	6	ADA21615	Ada21615 Human sec
31	1031	100.0	198	6	ADA10402	Ada10402 Human sec
32	1031	100.0	198	6	ADA17946	Ada17946 Human PRO
33	1031	100.0	198	6	ADA28054	Ada28054 Human sec
34	1031	100.0	198	6	ADA94634	Ada94634 Human sec
35	1031	100.0	198	6	ADA38859	Ada38859 Human sec
36	1031	100.0	198	6	ADA92980	Ada92980 Human sec
37	1031	100.0	198	7	ABO53237	Abu53237 Human sec
38	1031	100.0	198	7	ADA22541	Ada22541 Human sec
39	1031	100.0	198	7	ABO22607	Abu22607 Human sec
40	1031	100.0	198	7	ADA6707	Ada6707 Human sec
41	1031	100.0	198	7	ADA39400	Ada39400 Human sec
42	1031	100.0	198	7	ADB96426	Adb96426 Human PRO
43	1031	100.0	198	7	ADC57898	Adc57898 Human PRO
44	1031	100.0	198	7	ADC55262	Adc55262 Human PRO
45	1031	100.0	198	7	ADC12129	Adc12129 Human sec

ALIGNMENTS

RESULT 1
AAV88275
ID AAV88275 standard; protein; 198 AA.

AC AAV88275;

DT 16-OCT-2000 (first entry)

DE Human TANGO 184 protein.

XX TANGO 180; TANGO 181; TANGO 182; TANGO 183; TANGO 184; TANGO 185;

KW TANGO 186; TANGO 185; TANGO 189; TANGO 215; TANGO 187; human; murine;

XW secreted protein; transmembrane protein; gene therapy; vaccine;

KW diagnosis; treatment; detection.

OS Homo sapiens.

XX WO200018904-A2.

XX 06-APR-2000.

XX 30-SEP-1999; 99WO-US022817.

PR 30-SEP-1998; 98US-00164220.

PR 02-OCT-1998; 98US-00164169.

PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.

PI Barnes TM;

DR WPI: 2000-293144/25.

DR N-PSDB; AAA39941, AAA39942.

XX Isolated nucleic acids encoding TANGO polypeptides useful for preventing,

PT diagnosing and treating diseases associated with inappropriate protein

XX expression.

XX Claim 9; Fig 9; 249pp; English.

PS This invention describes novel human and murine nucleic acids encoding

CC TANGO polypeptides (which are either wholly secreted or transmembrane

CC proteins) which can be used for gene therapy and/or vaccination. The

CC peptides are designated TANGO 180 to TANGO 189 and TANGO 215. The nucleic

CC acids may be used to produce TANGO 180 to TANGO 189 and TANGO 215

CC polypeptides according to standard recombinant DNA methodologies. They

CC may also be used to detect and quantify the presence of TANGO nucleic

CC acids in a sample and therefore identify or diagnose diseases associated

CC with inappropriate TANGO expression (e.g. diseases related to over or

CC under expression of the polypeptides or the expression of inactive
CC polypeptides). The nucleic acids and the polypeptides they encode may be
CC used according to standard gene therapy protocols, to treat diseases
CC associated with inappropriate TANGO expression by supplementing a
CC patient's own production of the polypeptide or to rectify mutations that
CC may result in expression of an abnormally active polypeptide. The
CC polypeptides may also be used to identify and produce agonists and
CC antagonists of TANGO expression and activity which may be used to
CC modulate TANGO related processes and diseases. The polypeptides are
CC particularly useful for use as antigens for producing antibodies to TANGO
CC proteins which may be used for inhibiting the activity of TANGO proteins.
CC They may also be used to detect and quantify the presence of TANGO
CC proteins in samples and therefore identify patients in whom the protein
CC is over- or under-expressed. This sequence represents the human TANGO 184
CC protein described in the method of the invention

XX Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 3; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGGLRLGSLSLSCIALSVLLLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60
DB 1 MATLWGGLRLGSLSLSCIALSVLLLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60

QY 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCECKYEERSVTKVTIIYLSILGLLLYMV 120
DB 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCECKYEERSVTKVTIIYLSILGLLLYMV 120

QY 121 YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANAHVLAHSRANVINKVEYAQRWK 180
DB 121 YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANAHVLAHSRANVINKVEYAQRWK 180

QY 181 LQVQEQKSVFDRHVLS 198
DB 181 LQVQEQKSVFDRHVLS 198

RESULT 2

ID AAY66762
ID AAY66762 standard; protein; 198 AA.

AC AAY66762;

XX 05-APR-2000 (first entry)

XX Membrane-bound protein PRO1375.

XX Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
XX pharmaceutical; receptor immunoadhesin; gene mapping.

XX Homo sapiens.

XX WO9963088-A2.

XX 09-DEC-1999.

XX 02-JUN-1999; 99WO-US012252.

XX 02-JUN-1998; 98US-0087607P.

XX 02-JUN-1998; 98US-0087609P.

XX 02-JUN-1998; 98US-0087759P.

XX 03-JUN-1998; 98US-0087827P.

XX 04-JUN-1998; 98US-0088021P.

XX 04-JUN-1998; 98US-0088025P.

XX 04-JUN-1998; 98US-0088028P.

XX 04-JUN-1998; 98US-0088029P.

XX 04-JUN-1998; 98US-0088030P.

XX 04-JUN-1998; 98US-0088033P.

XX 05-JUN-1998; 98US-0088167P.

XX 05-JUN-1998; 98US-0088202P.

PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088730P.
PR 10-JUN-1998; 98US-0088734P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088741P.
PR 10-JUN-1998; 98US-0088742P.
PR 10-JUN-1998; 98US-0088810P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088858P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089440P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089532P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089600P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089801P.
PR 18-JUN-1998; 98US-0089907P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089947P.
PR 19-JUN-1998; 98US-0089948P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 23-JUN-1998; 98US-0090349P.
PR 23-JUN-1998; 98US-0090355P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090431P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090445P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090472P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090538P.
PR 24-JUN-1998; 98US-0090540P.
PR 24-JUN-1998; 98US-0090557P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090691P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 01-JUL-1998; 98US-0091358P.
PR 01-JUL-1998; 98US-0091360P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091519P.
PR 02-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091633P.
PR 02-JUL-1998; 98US-0091646P.
PR 02-JUL-1998; 98US-0091673P.

PR 07-JUL-1998; 98US-0091978P.
PR 07-JUL-1998; 98US-0091982P.
PR 09-JUL-1998; 98US-0092182P.
PR 10-JUL-1998; 98US-0092472P.
PR 20-JUL-1998; 98US-0093339P.
PR 30-JUL-1998; 98US-0094651P.
PR 04-AUG-1998; 98US-0095282P.
PR 04-AUG-1998; 98US-0095285P.
PR 04-AUG-1998; 98US-0095301P.
PR 04-AUG-1998; 98US-0095302P.
PR 04-AUG-1998; 98US-0095318P.
PR 04-AUG-1998; 98US-0095321P.
PR 04-AUG-1998; 98US-0095325P.
PR 04-AUG-1998; 98US-0095316P.
PR 10-AUG-1998; 98US-0095329P.
PR 10-AUG-1998; 98US-0096012P.
PR 10-AUG-1998; 98US-0096143P.
PR 11-AUG-1998; 98US-0096146P.
PR 12-AUG-1998; 98US-0096329P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096768P.
PR 17-AUG-1998; 98US-0096773P.
PR 17-AUG-1998; 98US-0096791P.
PR 17-AUG-1998; 98US-0096857P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096894P.
PR 17-AUG-1998; 98US-0096895P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096898P.
PR 18-AUG-1998; 98US-0096950P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0096960P.
PR 19-AUG-1998; 98US-0097022P.
PR 19-AUG-1998; 98US-0097141P.
PR 20-AUG-1998; 98US-0097218P.
PR 24-AUG-1998; 98US-0097661P.
PR 26-AUG-1998; 98US-0097951P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0097978P.
PR 26-AUG-1998; 98US-0097979P.
PR 26-AUG-1998; 98US-0097986P.
PR 26-AUG-1998; 98US-0098014P.
PR 31-AUG-1998; 98US-0098525P.
PR 16-SEP-1998; 98US-0100634P.
PR 12-JAN-1999; 99US-0115565P.
XX (GETH) GENENTECH INC.
XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
PI Wood WI, Yuan J;
XX WPI; 2000-072883/06.
DR N-PSDB; AAZ65108.
XX
PT Membrane-bound proteins and related nucleotide sequences.
XX
XX Claim 12; Fig 300; 822pp; English.
XX
XX The invention provides membrane-bound PRO polypeptides and
CC polynucleotides encoding them. The PRO sequences of the invention were
CC identified based on extracellular domain homology screening. The PRO
CC sequences have homology with proteins including LDL receptors, TIE
CC ligands and various enzymes. The membrane-bound proteins and receptor
CC molecules are useful as pharmaceutical and diagnostic agents. Receptor
CC immunoadhesins, for instance, can be used as therapeutic agents to block
CC receptor-ligand interactions. The membrane-bound proteins can also be
CC employed for screening of potential peptide or small molecule inhibitors
CC of the relevant receptor/ligand interaction. The PRO encoding sequences

are useful as hybridization probes, in chromosome and gene mapping and in
the generation of antisense RNA and DNA. PRO nucleic acid sequences will
also be useful for the preparation of PRO polypeptides, especially by
recombinant techniques
SQ Sequence 198 AA;
Query Match 100.0%; Score 1031; DB 3; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATLWGGLRLGSLLSCLALSLLAQLSDAAKPFEDVRCCKICPPYKENSHTYKN 60
DB 1 MATLWGGLRLGSLLSCLALSLLAQLSDAAKPFEDVRCCKICPPYKENSHTYKN 60
QY 61 ISQKDCCLHVPEMPVGRPDVEAYCLRCCKEYERSVTIKVTIIYLSILGLLLYV 120
DB 61 ISQKDCCLHVPEMPVGRPDVEAYCLRCCKEYERSVTIKVTIIYLSILGLLLYV 120
QY 121 YLTIVPEILKRLFGHAQLIQSDDDDIGDHOPFANAHDVLARSRRANVLKVEYAQORWK 180
DB 121 YLTIVPEILKRLFGHAQLIQSDDDDIGDHOPFANAHDVLARSRRANVLKVEYAQORWK 180
QY 181 LQVQEQKSVFDRHVLS 198
DB 181 LQVQEQKSVFDRHVLS 198
RESULT 3
AAV87231
ID AAY87231 standard; protein; 198 AA.
XX
AC AAY87231;
XX
DT 11-MAY-2000 (first entry)
XX
DE Human signal peptide containing protein HSP-8 SEQ ID NO:8.
XX
KW Human; signal peptide-containing protein; HSP; diagnosis; cancer;
KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;
KW antimicrobial; neurotropic; neuroprotective; cardiovascular; hepatotropic;
KW antiasthmatic; gene therapy; cell proliferation; neurological disorder;
KW reproductive disorder; developmental disorder; arteriosclerosis;
KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;
KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;
KW Parkinson's disease; Huntington's disease; ovulatory defect;
KW muscular dystrophy.
XX
OS Homo sapiens.
XX
PN WO200000610-A2.
XX
PD 06-JAN-2000.
XX
PF 25-JUN-1999; 99WO-US014484.
XX
PR 26-JUN-1998; 98US-0090762P.
PR 31-JUL-1998; 98US-0094983P.
PR 01-OCT-1998; 98US-0102686P.
PR 11-DEC-1998; 98US-0112129P.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn WR;
PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;
PI Bandman O;
XX
DR WPI; 2000-160673/14.
DR N-PSDB; AAZ98116.
XX
XX New human signal peptide-containing proteins useful in treatment,
PT prevention and diagnosis of e.g. cancer, inflammation and cardiovascular
PT disease.

XX PS Claim 1; Page 165; 327pp; English.

XX CC AAZ98109 to AAZ98242 encode AAY87224 to AAY87357 which represent the

XX CC human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have

XX CC anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatotropic,

XX CC neuroprotective, cardiovascular and antiasthmatic activities, and can be

XX CC used in gene therapy. HSPs can be used to treat or prevent disorders

XX CC associated with decreased activity or function of HSP. Antagonists of

XX CC HSP are used to treat or prevent disorders associated with increased

XX CC activity or function of HSP. Such diseases include cell proliferation

XX CC (including cancer), inflammation, cardiovascular, neurological,

XX CC reproductive or developmental disorders, (e.g. arteriosclerosis,

XX CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,

XX CC asthma, Crohn's disease, microbial or other infections, congestive or

XX CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's

XX CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSP

XX CC nucleic acids can be used for the recombinant production of HSP, for

XX CC detecting HSP in standard hybridisation and amplification assays (for

XX CC diagnosis and monitoring), in gene therapy, as antisense, triplex-forming

XX CC or ribozyme therapeutics, for detecting related sequences or genetic

XX CC variations, and for chromosomal mapping. HSP are also used to raise

XX CC specific antibodies (Ab) and to screen for agonists and antagonists

XX CC (potential therapeutic agents). Ab are used to diagnose, or monitor, HSP

XX CC -related diseases (in usual immunoassays), as therapeutic antagonists, in

XX CC competitive drug screens, and for purification of HSP from natural

XX CC sources

XX SQ Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 3; Length 198;

Best Local Similarity 100.0%; Pred. No. 1.4e-104;

Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGGLRLGSLLSCLALSLLVLLAQSLDAKNFEDVRCKICPPYKENSCHIYNKN 60

DB 1 MATLWGGLRLGSLLSCLALSLLVLLAQSLDAKNFEDVRCKICPPYKENSCHIYNKN 60

QY 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKEYERSSTVTKVTIIYLSILGLLLYV 120

DB 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKEYERSSTVTKVTIIYLSILGLLLYV 120

QY 121 YLTLPVLPILKRLFGHAQLIQSDDDIGDHPFANAHDLARSRSRANVLNKEVYAOQRWK 180

DB 121 YLTLPVLPILKRLFGHAQLIQSDDDIGDHPFANAHDLARSRSRANVLNKEVYAOQRWK 180

QY 181 LQVQQRKSVFDRHVLS 198

DB 181 LQVQQRKSVFDRHVLS 198

RESULT 4

AAV78807

ID AAY78807 standard; protein; 198 AA.

XX AC AAY78807;

XX DT 09-MAY-2000 (first entry)

XX DE Hydrophobic domain containing protein clone HP10529 protein sequence.

XX KW Hydrophobic domain; clone HP10529; nutritional supplement; SCID; HIV;

XX KW cell proliferation; immune stimulant; immune deficiency; tumour; pain;

XX KW rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;

XX KW myasthenia gravis; haematopoiesis regulator; tissue growth; depression;

XX KW anti-inflammatory; infection; bodily characteristic.

XX OS Homo sapiens.

XX PN WO200000506-A2.

XX PD 06-JAN-2000.

XX XX

PF 18-JUN-1999; 99WO-JP003242.

XX PR 26-JUN-1998; 98JP-00180008.

XX PA (SAGA) SAGAMI CHEM RES CENT.

XX PA (PROT-) PROTEGENE INC.

XX PI Kato S, Kimura T;

XX DR WPI; 2000-160665/14.

XX DR N-PSDB; AAZ90044, AAZ90054.

XX PT Novel human proteins having hydrophobic domains used for research and

XX PT diagnostic purposes.

XX PS Claim 1; Page 79-80; 117pp; English.

XX CC This sequence represents the hydrophobic domain containing protein, clone

XX CC HP10529 protein sequence. The sequence is isolated from a human

XX CC osteosarcoma cell line Soas-2. The invention relates to human proteins

XX CC with hydrophobic domains, the DNA and the cDNA encoding them. The

XX CC polynucleotides and proteins are predicted to have biological activities

XX CC which make them suitable for treating, preventing or ameliorating medical

XX CC conditions in humans and animals. Suggested activities include

XX CC nutritional activity (nutritional source or supplement); cytokine and

XX CC cell proliferation/differentiation activity; immune stimulating (e.g. as

XX CC vaccines) or suppressing activity (e.g. to treat various immune

XX CC deficiencies such as SCIDS or HIV, connective tissue disease, systemic

XX CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary

XX CC inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin

XX CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease

XX CC and autoimmune inflammatory eye disease, as well as asthma, allergies and

XX CC organ transplantation; haematopoiesis regulating activity (e.g. in

XX CC treatment of myeloid or lymphoid cell deficiencies); tissue growth

XX CC activity (e.g. wound healing and tissue repair, ulcers, burns,

XX CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic

XX CC activity; haemostatic and thrombolytic activity (e.g. treating

XX CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and

XX CC tumour inhibition activity. The polynucleotides are also stated to be

XX CC useful for gene therapy. Other activities include inhibiting infections

XX CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,

XX CC malaria); effecting bodily characteristics such as, e.g. weight, colour,

XX CC skin, effecting biorhythms or cardiac cycles; enhancing fertility;

XX CC treatment of depression; treatment of pain; hormonal or endocrine

XX CC activity. The polynucleotides may also be used for recombinant expression

XX CC of the protein

XX SQ Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 3; Length 198;

Best Local Similarity 100.0%; Pred. No. 1.4e-104;

Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGGLRLGSLLSCLALSLLVLLAQSLDAKNFEDVRCKICPPYKENSCHIYNKN 60

DB 1 MATLWGGLRLGSLLSCLALSLLVLLAQSLDAKNFEDVRCKICPPYKENSCHIYNKN 60

QY 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKEYERSSTVTKVTIIYLSILGLLLYV 120

DB 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKEYERSSTVTKVTIIYLSILGLLLYV 120

QY 121 YLTLPVLPILKRLFGHAQLIQSDDDIGDHPFANAHDLARSRSRANVLNKEVYAOQRWK 180

DB 121 YLTLPVLPILKRLFGHAQLIQSDDDIGDHPFANAHDLARSRSRANVLNKEVYAOQRWK 180

QY 181 LQVQQRKSVFDRHVLS 198

DB 181 LQVQQRKSVFDRHVLS 198

RESULT 5

AAV93740

ID AAM93740 standard; protein; 198 AA.

QY 1 MATWGLRLGSLLSCLALSLLAQLSDRAKNFEDVRCICPPYKENSCHYNNK 60
DB 1 MATWGLRLGSLLSCLALSLLAQLSDRAKNFEDVRCICPPYKENSCHYNNK 60
QY 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKEYERSVTKVTIIYLSILGLLLYMW 120
DB 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKEYERSVTKVTIIYLSILGLLLYMW 120
QY 121 YLTVPEPLKRLFGHAQLIQSDDDIGDHPFANAHDVLAARSRAVNLKVEYAQRWK 180
DB 121 YLTVPEPLKRLFGHAQLIQSDDDIGDHPFANAHDVLAARSRAVNLKVEYAQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
DB 181 LQVQQRKSVFDRHVLS 198
RESULT 7
ID AAB20120
XX AAB20120 standard; protein; 198 AA.
AC AAB20120;
DT 30-APR-2001 (first entry)
DE Human immunostimulant PRO1375.
KW PRO1375; UNQ712; human; immune disease; autoimmune disease;
KW antirheumatic; antiarthritic; antiinflammatory; antianaemic;
KW immunosuppressive; antithyroid; antidiabetic; neuroprotective;
KW hepatotropic; virucide; dermatological; antipsoriatic; antiasthmatic;
KW antiallergic; immunostimulant.
XX
XX Homo sapiens.
FH Key Location/Qualifiers
FT Domain 11..28
FT /note= "transmembrane domain type II"
FT Modified-site 12..18
FT /note= "N-myristoylation site"
FT Modified-site 60..64
FT /note= "Asn is N-glycosylated"
FT Modified-site 78..86
FT /note= "tyrosine kinase phosphorylation site"
FT Domain 103..125
FT /note= "transmembrane domain"
XX
XX WO200105972-A1.
XX
XX 25-JAN-2001.
XX
XX 15-MAR-2000; 2000WO-US006884.
XX
XX 20-JUL-1999; 99US-0144758P.
XX
XX (GETH) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Fong S, Goddard A, Godowski PJ;
PI Gurney AL, Hillan KJ, Mark MR, Marsters SA, Pitti RM, Tumas D;
PI Watanabe CK, Wood WI;
XX WPI; 2001-103149/11.
DR N-PSDB; AAF30062.
XX
XX New PRO polypeptides, nucleic acids and (ant)agonists, useful for
PT diagnosing and treating immune-related disorders, such as multiple
PT sclerosis, rheumatoid arthritis and diabetes.
XX
XX Claim 20; Fig 26; 127pp; English.
XX
XX The present sequence is that of novel human immunomodulator PRO1375
CC (UNQ712), as deduced from cDNA (see AAF30062) isolated from a human

CC pancreas tissue library. PRO1375 has a mol.wt. of 22 kDa and a pI of
CC 8.47. The invention provides polynucleotides (see AAF30050-62) encoding
CC novel human PRO proteins (see AAB20108-20) including PRO1375. Claimed
CC compositions comprising these proteins or their agonists are useful for
CC increasing infiltration of inflammatory cells into a tissue of a mammal,
CC stimulating or enhancing an immune response, or increasing the
CC proliferation of T-lymphocytes in a mammal in response to an antigen.
CC Claimed compositions comprising a PRO polypeptide or its antagonist have
CC the opposite effect. A claimed method for creating an immune related
CC disorder, such as a T cell disorder, involves administering a PRO
CC polypeptide, an agonist antibody or an antagonist antibody. The disorder
CC is selected from systemic lupus erythematosus, rheumatoid arthritis,
CC osteoarthritis, juvenile chronic arthritis, spondyloarthropathy, systemic
CC sclerosis, idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune renal
CC thrombocytopaenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, demyelinated diseases (such as multiple sclerosis), autoimmune
CC chronic active hepatitis, primary biliary cirrhosis, granulomatous
CC hepatitis, sclerosing cholangitis, inflammatory bowel disease (ulcerative
CC colitis and Crohn's disease), gluten-sensitive enteropathy, Whipple's
CC disease, (auto)immune-mediated skin diseases (such as bullous skin
CC disease, erythema multiforme and psoriasis), allergic diseases (such as
CC asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity and
CC urticaria), immunologic diseases of the lung and transplantation
CC associated diseases (such as graft rejection and graft-versus-host
CC disease) (all claimed). Claimed methods of diagnosing these disorders
CC comprise detecting the level of expression of the PRO gene. Also claimed
CC are a method of identifying a compound capable of inhibiting the
CC expression or activity of the PRO polypeptide, vectors, host cells,
CC antibodies, and a method of stimulating the proliferation of T
CC lymphocytes using PRO1375
XX
SQ Sequence 198 AA;
Query Match 100.0%; Score 1031; DB 4; Length 198;
Best Local Similarity 100.0%; Pred No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATWGLRLGSLLSCLALSLLAQLSDAANKFEDVRCICPPYKENSCHYNNK 60
DB 1 MATWGLRLGSLLSCLALSLLAQLSDAANKFEDVRCICPPYKENSCHYNNK 60
QY 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKEYERSVTKVTIIYLSILGLLLYMW 120
DB 61 ISQKDCDCLHVPMPVRGPDVEAYCLRCCKEYERSVTKVTIIYLSILGLLLYMW 120
QY 121 YLTVPEPLKRLFGHAQLIQSDDDIGDHPFANAHDVLAARSRAVNLKVEYAQRWK 180
DB 121 YLTVPEPLKRLFGHAQLIQSDDDIGDHPFANAHDVLAARSRAVNLKVEYAQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
DB 181 LQVQQRKSVFDRHVLS 198
RESULT 8
ID AAM38735
XX AAM38735 standard; protein; 198 AA.
AC AAM38735;
DT 22-OCT-2001 (first entry)
DE Human polypeptide SEQ ID NO 1880.
XX
XX Human; notropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.

```
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US034263.
XX PR 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-00552317.
XX PR 20-JUN-2000; 2000US-00598042.
XX PR 19-JUL-2000; 2000US-00620312.
XX PR 03-AUG-2000; 2000US-00653450.
XX PR 14-SEP-2000; 2000US-00662191.
XX PR 19-OCT-2000; 2000US-00693036.
XX PR 29-NOV-2000; 2000US-00727344.
XX PA (HYSE-) HYSEQ INC.
XX PF Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX PI Zhou P, Goodrich R, Drmanac RT;
XX DR WPI; 2001-442253/47.
XX DR N-PSDB; AAI57891.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders such
XX PI as central nervous system injuries.
XX PS Example 3; SEQ ID NO 1880; 10078pp; English.
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
XX CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: Immune system suppression,
XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC assays for receptor activity, arthritis and inflammation, leukaemias and
XX CC C.N.S disorders. Note: The sequence data for this patent did not form
XX CC part of the printed specification
XX SQ Sequence 198 AA;
Query Match 100.0%; Score 1031; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104; Mismatches 0; Gaps 0;
Matches 198; Conservative 0; Indels 0;
QY 1 MATLWGLRLGSLSLALSLLLAQLSDAKNFEDVRCKICPPYKENSIGHYKN 60
Db 1 MATLWGLRLGSLSLALSLLLAQLSDAKNFEDVRCKICPPYKENSIGHYKN 60
QY 61 ISQKDCDCLHVPVPGDPVEAYCLRCCKYEERSVTIKVIIIVLSILGLLLYV 120
Db 61 ISQKDCDCLHVPVPGDPVEAYCLRCCKYEERSVTIKVIIIVLSILGLLLYV 120
QY 121 YLTLPVLPKRLFGHQLIQSDDDIGDHPQFANAHDLARSGRANVLNVEYAQRWK 180
Db 121 YLTLPVLPKRLFGHQLIQSDDDIGDHPQFANAHDLARSGRANVLNVEYAQRWK 180
QY 181 LQVQQRKSVFDRHVLS 198
Db 181 LQVQQRKSVFDRHVLS 198
RESULT 9
AAB50926
ID AAB50926 standard; protein; 198 AA.
```

```
XX AC AAB50926;
XX DT 21-MAR-2001 (first entry)
XX DE Human PRO1375 protein.
XX KW Human; PRO; antiinflammatory; dermatological; antiarthritic;
XX KW antirheumatic; cardiac; antianemic; immunosuppressive; antithyroid;
XX KW antidiabetic; neuroprotective; hepatotropic; virucide;
XX KW antiallergic; antiasthmatic; immune related disorder;
XX KW hepatobiliary disease; autoimmune disease; allergy.
XX OS Homo sapiens.
XX PN WO200073452-A2.
XX PD 07-DEC-2000.
XX PF 02-JUN-2000; 2000WO-US015264.
XX PR 02-JUN-1999; 99WO-US012252.
XX PR 20-JUL-1999; 99US-0144732P.
XX PR 20-JUL-1999; 99US-0144758P.
XX PR 28-JUL-1999; 99US-0146222P.
XX PR 01-SEP-1999; 99WO-US020111.
XX PR 15-SEP-1999; 99WO-US021090.
XX PR 15-SEP-1999; 99WO-US021547.
XX PR 29-OCT-1999; 99US-0162506P.
XX PR 30-NOV-1999; 99WO-US028313.
XX PR 01-DEC-1999; 99WO-US028634.
XX PR 02-DEC-1999; 99WO-US028551.
XX PR 09-DEC-1999; 99WO-US028565.
XX PR 20-DEC-1999; 99US-0170262P.
XX PR 05-JAN-2000; 2000WO-US000219.
XX PR 06-JAN-2000; 2000WO-US000376.
XX PR 11-FEB-2000; 2000WO-US003565.
XX PR 18-FEB-2000; 2000WO-US004341.
XX PR 22-FEB-2000; 2000WO-US004414.
XX PR 24-FEB-2000; 2000WO-US004914.
XX PR 24-FEB-2000; 2000WO-US005004.
XX PR 01-MAR-2000; 2000WO-US005601.
XX PR 02-MAR-2000; 2000WO-US005841.
XX PR 03-MAR-2000; 2000US-0187202P.
XX PR 15-MAR-2000; 2000WO-US006884.
XX PR 20-MAR-2000; 2000WO-US007377.
XX PR 21-MAR-2000; 2000WO-US007532.
XX PR 30-MAR-2000; 2000WO-US008439.
XX PR 17-MAY-2000; 2000WO-US013705.
XX PR 22-MAY-2000; 2000WO-US014042.
XX PF (GETH ) GENENTECH INC.
XX PI Ashkenazi AJ, Baker KP, Chan B, Goddard A, Godowski PJ;
XX PI Gurney AL, Hebert C, Henzel W, Kabakoff RC, Shelton DL, Tumas D;
XX PI Watanabe CK, Wood WI;
XX DR WPI; 2001-025253/03.
XX DR N-PSDB; AAC91485.
XX PT Thirty three nucleic acids encoding PRO polypeptides which are useful in
XX PT the diagnosis and treatment of immune related disorders, e.g. systemic
XX PT lupus erythematosus, rheumatoid arthritis, osteoarthritis, thyroiditis
XX PT and diabetes mellitus.
XX PS Claim 58; Fig 50; 218pp; English.
XX CC The present sequence is one of thirty three novel PRO polypeptides. The
XX CC PRO polypeptides, anti-PRO antibodies, agonists and antagonists are
XX CC useful for treating and diagnosing immune related disorders such as
XX CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
```

CC juvenile chronic arthritis, spondyloarthropathies, systemic sclerosis,
 CC idiopathic inflammatory myopathies, Sjogren's syndrome, systemic
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 CC disease, demyelinating diseases of the central and peripheral nervous
 CC systems (such as multiple sclerosis, idiopathic demyelinating
 CC polyneuropathy or Guillain-Barre syndrome, and chronic inflammatory
 CC demyelinating polyneuropathy), hepatobiliary diseases (such as
 CC infectious, granulomatous chronic active hepatitis, primary biliary
 CC cirrhosis, granulomatous hepatitis and sclerosing cholangitis)
 CC inflammatory bowel disease, gluten-sensitive enteropathy and Whipple's
 CC disease, autoimmune or immune-mediated skin diseases (such as bullous
 CC skin diseases, erythema multiforme, contact dermatitis, psoriasis),
 CC allergic diseases such as asthma, allergic rhinitis, atopic dermatitis,
 CC food hypersensitivity and urticaria), immunological diseases of the lung
 CC (such as eosinophilic pneumonias, idiopathic pulmonary fibrosis and
 CC hypersensitivity pneumonitis), transplantation associated diseases
 CC including graft rejection and graft-versus-host diseases
 XX
 SQ Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 4; Length 198;
 Best Local Similarity 100.0%; Pred. No. 1.4e-104;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGGLRLGSLSLSCIALSVLLLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60
 DB 1 MATLWGGLRLGSLSLSCIALSVLLLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60
 QY 61 ISQKDCCLHVEPMPVGRPDVEAYCLURCECKYEERSSVTIKVTIIYLSILGLLLYMW 120
 DB 61 ISQKDCCLHVEPMPVGRPDVEAYCLURCECKYEERSSVTIKVTIIYLSILGLLLYMW 120
 QY 121 YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANADVLRSSRANVLNKVEYAQQRWK 180
 DB 121 YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANADVLRSSRANVLNKVEYAQQRWK 180
 QY 181 LQVQEQKSVFDRHVLS 198
 DB 181 LQVQEQKSVFDRHVLS 198

RESULT 10
 AAB65285
 ID AAB65285 standard; protein; 198 AA.
 AC AAB65285;
 XX
 XX
 DT 02-APR-2001 (first entry)
 XX
 DE Human PRO1375 (UNO712) protein sequence SEQ ID NO:418.
 XX
 KW Human; secreted and transmembrane protein; PRO; cytostatic; cell death;
 KW cancer; chromosomal mapping; Gene mapping; tissue typing;
 KW diagnostic assay.
 XX
 OS Homo sapiens.
 XX
 PN WO2000073454-A1.
 XX
 PD 07-DEC-2000.
 XX
 PF 30-MAR-2000; 2000WO-US008439.
 XX
 PR 02-JUN-1999; 99WO-US012252.
 PR 23-JUN-1999; 99US-0141037P.
 PR 07-JUL-1999; 99US-0143048P.
 PR 20-JUL-1999; 99US-0144758P.
 PR 26-JUL-1999; 99US-0145698P.
 PR 28-JUL-1999; 99US-0146222P.
 PR 17-AUG-1999; 99US-0149366P.
 PR 15-SEP-1999; 99WO-US021090.
 PR 15-SEP-1999; 99WO-US021547.

PR 08-OCT-1999; 99US-0158663P.
 PR 30-NOV-1999; 99WO-US028313.
 PR 01-DEC-1999; 99WO-US028301.
 PR 16-DEC-1999; 99WO-US030095.
 PR 20-DEC-1999; 99WO-US030911.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 28-FEB-2000; 2000WO-US005004.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 20-MAR-2000; 2000WO-US007377.
 XX
 PA (CBTH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
 PI Ferrara N, Fong S, Gerber H, Gerritsen MP, Goddard A, Godowski PJ;
 PI Grimaldi CJ, Turney AL, Kijavini IJ, Napier MA, Pan J, Paoni NF;
 PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
 PI Zhang Z;
 XX
 DR WPI; 2001-032160/04.
 N-PSDB; AAF44254.
 XX
 CC

PRO polynucleotides used to produce polypeptides used to target bioactive
 molecules such as toxins, radiolabels or antibodies, to specific cells,
 to cause targeted cell death.
 Claim 12; Fig 300; 935pp; English.

The present invention describes human secreted and transmembrane PRO
 proteins. The PRO proteins have cytostatic activity. The PRO proteins can
 be used for targeted delivery of bioactive molecules, such as toxins,
 radiolabels or antibodies, that cause cell death. PRO nucleotide
 sequences, and their fragments, can be used as hybridisation probes, in
 chromosomal and gene mapping, and in the generation of anti-sense RNA and
 DNA. They may also be used to produce transgenic animals which are used
 to develop and screen therapeutically useful reagents. The PRO nucleotide
 and protein sequence can be used for tissue typing and in treating
 cancer. Anti-PRO antibodies can be used in diagnostic assays. AAF44270 to
 AAF44470 represent PCR primers and hybridisation probes used in the
 isolation of human PRO sequences. AAF44097 to AAF44269 and AAB65154 to
 AAB65300 represent human PRO polynucleotide and protein sequences given
 in the exemplification of the present invention

Sequence 198 AA;
 Query Match 100.0%; Score 1031; DB 4; Length 198;
 Best Local Similarity 100.0%; Pred. No. 1.4e-104;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATLWGGLRLGSLSLSCIALSVLLLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60
 DB 1 MATLWGGLRLGSLSLSCIALSVLLLAQLSDAAKNFEDVRCKICPPYKENSCHIYKN 60
 QY 61 ISQKDCCLHVEPMPVGRPDVEAYCLURCECKYEERSSVTIKVTIIYLSILGLLLYMW 120
 DB 61 ISQKDCCLHVEPMPVGRPDVEAYCLURCECKYEERSSVTIKVTIIYLSILGLLLYMW 120
 QY 121 YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANADVLRSSRANVLNKVEYAQQRWK 180
 DB 121 YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANADVLRSSRANVLNKVEYAQQRWK 180
 QY 181 LQVQEQKSVFDRHVLS 198
 DB 181 LQVQEQKSVFDRHVLS 198

RESULT 11
 AAB61428


```
PR 19-JUN-1998; 98US-00899552P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 26-AUG-1998; 98US-0090254P.
PR 23-JUN-1998; 98US-0090349P.
PR 23-JUN-1998; 98US-0090355P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090431P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090445P.
PR 24-JUN-1998; 98US-0090472P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 24-JUN-1998; 98US-0090542P.
PR 24-JUN-1998; 98US-0090577P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-0090862P.
PR 01-JUL-1998; 98US-0091360P.
PR 01-JUL-1998; 98US-0091344P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091519P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091633P.
PR 02-JUL-1998; 98US-0091646P.
PR 02-JUL-1998; 98US-0091673P.
PR 07-JUL-1998; 98US-0091578P.
PR 07-JUL-1998; 98US-0091982P.
PR 09-JUL-1998; 98US-0092182P.
PR 10-JUL-1998; 98US-0092472P.
PR 20-JUL-1998; 98US-0093339P.
PR 30-JUL-1998; 98US-0094651P.
PR 04-AUG-1998; 98US-0095282P.
PR 04-AUG-1998; 98US-0095285P.
PR 04-AUG-1998; 98US-0095301P.
PR 04-AUG-1998; 98US-0095302P.
PR 04-AUG-1998; 98US-0095318P.
PR 04-AUG-1998; 98US-0095321P.
PR 04-AUG-1998; 98US-0095325P.
PR 10-AUG-1998; 98US-0095929P.
PR 10-AUG-1998; 98US-0096012P.
PR 11-AUG-1998; 98US-0096143P.
PR 11-AUG-1998; 98US-0096446P.
PR 12-AUG-1998; 98US-0096329P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096768P.
PR 17-AUG-1998; 98US-0096773P.
PR 17-AUG-1998; 98US-0096791P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096895P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096950P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0096960P.
PR 18-AUG-1998; 98US-0097022P.
PR 19-AUG-1998; 98US-0097141P.
PR 20-AUG-1998; 98US-0097218P.
PR 24-AUG-1998; 98US-0097661P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.

PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0097978P.
PR 26-AUG-1998; 98US-0097979P.
PR 26-AUG-1998; 98US-0097986P.
PR 26-AUG-1998; 98US-0098014P.
PR 21-AUG-1998; 98US-0098525P.
PR 16-SEP-1998; 98US-0100634P.
PR 16-SEP-1998; 98US-0100634P.
PR 17-SEP-1998; 98US-0100858P.
PR 17-SEP-1998; 98US-0100858P.
PR 07-OCT-1998; 98US-0100858P.
PR 01-DEC-1998; 98US-0100858P.
PR 22-DEC-1998; 98US-0113296P.
PR 05-JAN-1999; 98US-0113296P.
PR 08-MAR-1999; 98US-0113296P.
PR 12-MAR-1999; 98US-0123957P.
PR 12-MAR-1999; 98US-0123957P.
PR 23-JUN-1999; 98US-0141037P.
PR 07-JUL-1999; 98US-0143048P.
PR 26-JUL-1999; 98US-0144758P.
PR 26-JUL-1999; 98US-0145698P.
PR 28-JUL-1999; 98US-0146222P.
PR 17-AUG-1999; 98US-0149396P.
PR 15-SEP-1999; 98US-0201090P.
PR 15-SEP-1999; 98US-0201547P.
PR 08-OCT-1999; 98US-0158663P.
PR 30-NOV-1999; 98US-0158663P.
PR 01-DEC-1999; 98US-0202831P.
PR 16-DEC-1999; 98US-0202863P.
PR 16-DEC-1999; 98US-0203009P.
PR 20-DEC-1999; 98US-0203091P.
PR 05-JAN-2000; 2000US-0200021P.
PR 06-JAN-2000; 2000US-0200037P.
PR 11-FEB-2000; 2000US-0200037P.
PR 18-FEB-2000; 2000US-0200037P.
PR 22-FEB-2000; 2000US-0200041P.
PR 24-FEB-2000; 2000US-0200041P.
PR 24-FEB-2000; 2000US-0200041P.
PR 24-FEB-2000; 2000US-0200041P.
PR 02-MAR-2000; 2000US-0200050P.
PR 02-MAR-2000; 2000US-0200050P.
PR 10-MAR-2000; 2000US-0200063P.
PR 13-MAR-2000; 2000US-0200068P.
PR 20-MAR-2000; 2000US-0200073P.
PR 30-MAR-2000; 2000US-0200084P.
PR 15-MAY-2000; 2000US-0201335P.
PR 17-MAY-2000; 2000US-0201370P.
PR 22-MAY-2000; 2000US-0201404P.
PR 30-MAY-2000; 2000US-0201494P.
PR 02-JUN-2000; 2000US-0201524P.
PR 23-JUN-2000; 2000US-0201563P.
PR 28-JUL-2000; 2000US-0202071P.
PR 11-AUG-2000; 2000US-0202031P.
PR 23-AUG-2000; 2000US-0202322P.
PR 24-AUG-2000; 2000US-0202328P.
PR 07-SEP-2000; 2000US-0203097P.

Query Match 100.0%; Score 1031; DB 6; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATLWGLRLGSLLSLCLSLALSVLLAQLSDAAKNFEDVCKICPPYKENGSHIYNKN 60
Db 1 MATLWGLRLGSLLSLCLSLALSVLLAQLSDAAKNFEDVCKICPPYKENGSHIYNKN 60
Qy 61 ISQKDCDCLHVVEPVPVGPVVEAYCLFCECKYBERSSTIKVTIITVLSILGLLLYV 120
Db 61 ISQKDCDCLHVVEPVPVGPVVEAYCLFCECKYBERSSTIKVTIITVLSILGLLLYV 120
Qy 121 YLTLVEPIKRLFGHAQLIQSDDDIGHQPFANAHVLAERSRANVKNKVEVAQRWK 180
Db 121 YLTLVEPIKRLFGHAQLIQSDDDIGHQPFANAHVLAERSRANVKNKVEVAQRWK 180
Qy 181 LQVQEQKSVFDRHVLS 198
```

Db 181 LQVQEQKSVFDRHVLS 198
 RESULT 13
 ABUS9178
 ID ABUS9178 standard; protein; 198 AA.
 XX AC ABUS9178;
 XX DT 28-APR-2003 (first entry)
 XX DE Novel human secreted or transmembrane protein PRO1375.
 XX KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
 KW cardiac insufficiency disorder; cancer; tumour; immune response;
 KW adrenal cortical capillary endothelial growth; c-fos induction;
 KW vascular endothelial growth factor inhibition; VEGF inhibition;
 KW endothelial cell growth inhibitor; T-lymphocytes stimulation;
 KW retinal neurons cell survival; rod photoreceptor cell survival;
 KW mammalian kidney mesangial cell proliferation; kidney disorder;
 KW mammary gland epithelial cell proliferation; Berger disease;
 KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;
 KW chondrocyte redifferentiation; sports injury; arthritis.
 XX OS Homo sapiens.
 XX PN US2002132252-A1.
 XX PD 19-SEP-2002.
 XX PF 14-NOV-2001; 2001US-00990442.
 PR 16-JUN-1997; 97US-0049787P.
 PR 17-OCT-1997; 97US-0062250P.
 PR 03-NOV-1997; 97WO-US020069.
 PR 12-NOV-1997; 97US-0065186P.
 PR 13-NOV-1997; 97US-0065311P.
 PR 24-NOV-1997; 97US-0066770P.
 PR 25-FEB-1998; 98US-0075945P.
 PR 20-MAR-1998; 98US-0078910P.
 PR 28-APR-1998; 98US-0083322P.
 PR 07-MAY-1998; 98US-0084600P.
 PR 28-MAY-1998; 98US-0087106P.
 PR 02-JUN-1998; 98US-0087607P.
 PR 02-JUN-1998; 98US-0087609P.
 PR 02-JUN-1998; 98US-0087759P.
 PR 03-JUN-1998; 98US-0087827P.
 PR 04-JUN-1998; 98US-0088021P.
 PR 04-JUN-1998; 98US-0088025P.
 PR 04-JUN-1998; 98US-0088026P.
 PR 04-JUN-1998; 98US-0088028P.
 PR 04-JUN-1998; 98US-0088029P.
 PR 04-JUN-1998; 98US-0088030P.
 PR 04-JUN-1998; 98US-0088033P.
 PR 04-JUN-1998; 98US-0088326P.
 PR 05-JUN-1998; 98US-0088167P.
 PR 05-JUN-1998; 98US-0088202P.
 PR 05-JUN-1998; 98US-0088212P.
 PR 05-JUN-1998; 98US-0088217P.
 PR 09-JUN-1998; 98US-0088655P.
 PR 10-JUN-1998; 98US-0088734P.
 PR 10-JUN-1998; 98US-0088738P.
 PR 10-JUN-1998; 98US-0088742P.
 PR 10-JUN-1998; 98US-0088810P.
 PR 10-JUN-1998; 98US-0088824P.
 PR 10-JUN-1998; 98US-0088826P.
 PR 11-JUN-1998; 98US-0088858P.
 PR 11-JUN-1998; 98US-0088861P.
 PR 11-JUN-1998; 98US-0088876P.
 PR 12-JUN-1998; 98US-0089105P.
 PR 16-JUN-1998; 98US-0089440P.
 PR 16-JUN-1998; 98US-0089512P.
 PR 16-JUN-1998; 98US-0089514P.
 PR 17-JUN-1998; 98US-0089532P.
 PR 17-JUN-1998; 98US-0089538P.
 PR 17-JUN-1998; 98US-0089598P.
 PR 17-JUN-1998; 98US-0089599P.
 PR 17-JUN-1998; 98US-0089600P.
 PR 17-JUN-1998; 98US-0089653P.
 PR 18-JUN-1998; 98US-0089801P.
 PR 18-JUN-1998; 98US-0089907P.
 PR 18-JUN-1998; 98US-0089908P.
 PR 16-SEP-1998; 98WO-US019330.
 PR 17-SEP-1998; 98WO-US019437.
 PR 07-OCT-1998; 98WO-US021141.
 PR 01-DEC-1998; 98WO-US025108.
 PR 05-JAN-1999; 99WO-US000106.
 PR 08-MAR-1999; 99WO-US005028.
 PR 02-JUN-1999; 99WO-US012252.
 PR 15-SEP-1999; 99WO-US021090.
 PR 15-SEP-1999; 99WO-US021547.
 PR 30-NOV-1999; 99WO-US028313.
 PR 01-DEC-1999; 99WO-US028301.
 PR 01-DEC-1999; 99WO-US028634.
 PR 16-DEC-1999; 99WO-US030095.
 PR 20-DEC-1999; 99WO-US030911.
 PR 06-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003585.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 10-MAR-2000; 2000WO-US006319.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 20-MAR-2000; 2000WO-US007377.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 15-MAY-2000; 2000WO-US013358.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 11-AUG-2000; 2000WO-US022031.
 PR 23-AUG-2000; 2000WO-US023522.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 28-FEB-2001; 2001WO-US006520.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR 28-AUG-2001; 2001US-00941992.
 (GETH) GENENTECH INC.
 PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL,
 PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ,
 PI Grimaldi JC, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF,
 PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams FM, Wood WI,
 PI Zhang Z;
 XX WPI; 2003-247083/24.
 XX N-PSDB; ABX80390.
 XX Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346
 XX and PRO1375, which stimulate proliferation of stimulated T-lymphocytes
 XX are therapeutically useful for enhancing immune response and in cancer
 XX treatments.
 XX Claim 12; Fig 300; 648pp; English.
 XX The invention describes an isolated human PRO polypeptide. The PRO

polypeptides are useful in detecting PRO polypeptides in a sample, in linking a bioactive molecule to a cell expressing a PRO polypeptide, and in modulating at least one biological activity of a cell expressing a PRO polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186 stimulate adrenal cortical capillary endothelial growth, and PRO536, PRO943, PRO828, PRO826, PRO1068 or PRO535, PRO826, PRO819, PRO1126, PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus useful for treating conditions or disorders where angiogenesis would be beneficial, e.g. wound healing and antagonist of this polypeptide are useful for treating cancerous tumours. PRO812 inhibits vascular endothelial growth factor (VEGF) stimulated proliferation of endothelial cells and is thus useful for inhibiting endothelial cell growth in mammals which would be beneficial in inhibiting tumour growth. PRO826, PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of stimulated T-lymphocytes and are therapeutically useful for enhancing immune response. PRO828, PRO826, PRO1068 or PRO1132 enhance survival of retinal neurons cells (PRO1132 is also enhances survival/proliferation of rod photoreceptor cells) and therefore are useful for treating retinal disorders of injuries, e.g. retinitis pigmentosa, AMD. PRO819, PRO813 and PRO1066 induce proliferation of mammalian kidney mesangial cells, and therefore are useful for treating kidney disorders associated with decreased mesangial cell function such as Berger disease or other nephropathies associated with dermatitis, herpeticiformis or Crohn's disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the proliferation and/or redifferentiation of chondrocytes in culture and are thus useful for treating sports injuries, and arthritis. This is the amino acid sequence of a novel human PRO protein

XX Sequence 198 AA;

Query Match 100.0%; Score 1031; DB 6; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.4e-104;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGLRLGSLLSCLALSLLAQLSDAANKFEDVRCKICPPYKENSCHIYNN 60
Db 1 MATLWGLRLGSLLSCLALSLLAQLSDAANKFEDVRCKICPPYKENSCHIYNN 60

QY 61 ISQKDCCLHVPEMPVGPDPVEAYCLRCCKEYERSSVTIKTIIIVLSILGLLLYV 120
Db 61 ISQKDCCLHVPEMPVGPDPVEAYCLRCCKEYERSSVTIKTIIIVLSILGLLLYV 120

QY 121 YLTLPETLKRFLFGHAQLIQSDDDIGDHQFFANAHVDVLARGSRANVLNKVEYAQRWK 180
Db 121 YLTLPETLKRFLFGHAQLIQSDDDIGDHQFFANAHVDVLARGSRANVLNKVEYAQRWK 180

QY 181 LQVQQRKSVFDRHWLS 198
Db 181 LQVQQRKSVFDRHWLS 198

RESULT 14
ID ABU82690 standard; protein; 198 AA.
XX AC ABU82690;
XX DT 26-JUN-2003 (first entry)
XX DE Human secreted/transmembrane protein PRO1375.
XX KW Human; PRO; secreted protein; transmembrane protein;
KW cardiac insufficiency disorders; angiogenesis; wound healing;
KW cancerous tumour; immune response; retinal disorder; sight loss;
KW retinitis pigmentosa; age-related macular degeneration; AMD;
KW kidney disorder; Berger disease; nephropathy; dermatitis; herpeticiformis;
KW Crohn's disease; sports injury; arthritis.
XX OS Homo sapiens.
XX PN US2003032023-A1.
XX

PD 13-FEB-2003.
XX 14-NOV-2001; 2001US-00990711.
XX 16-JUN-1997; 97US-0049787P.
PR 17-OCT-1997; 97US-0062250P.
PR 05-NOV-1997; 97WO-US020069.
PR 12-NOV-1997; 97US-0061886P.
PR 13-NOV-1997; 97US-0065311P.
PR 24-NOV-1997; 97US-0066770P.
PR 25-FEB-1998; 98US-0075945P.
PR 20-MAR-1998; 98US-0078910P.
PR 28-APR-1998; 98US-0083323P.
PR 07-MAY-1998; 98US-0084600P.
PR 28-MAY-1998; 98US-0087106P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088021P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088026P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088030P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088734P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088742P.
PR 10-JUN-1998; 98US-0088810P.
PR 10-JUN-1998; 98US-0088824P.
PR 11-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088858P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089440P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089532P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089599P.
PR 17-JUN-1998; 98US-0089600P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089601P.
PR 18-JUN-1998; 98US-0089907P.
PR 19-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089947P.
PR 19-JUN-1998; 98US-0089948P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 23-JUN-1998; 98US-0090349P.
PR 23-JUN-1998; 98US-0090355P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090431P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090443P.
PR 24-JUN-1998; 98US-0090445P.
PR 24-JUN-1998; 98US-0090472P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 24-JUN-1998; 98US-0090542P.
PR 25-JUN-1998; 98US-0090577P.
PR 25-JUN-1998; 98US-0090676P.

PR	25-JUN-1998;	98US-0090678P.	PR	02-JUN-1999;	99WO-US012252.
PR	25-JUN-1998;	98US-0090690P.	PR	23-JUN-1999;	99US-0141037P.
PR	25-JUN-1998;	98US-0090694P.	PR	07-JUL-1999;	99US-0143048P.
PR	25-JUN-1998;	98US-0090695P.	PR	20-JUL-1999;	99US-0144758P.
PR	25-JUN-1998;	98US-0090696P.	PR	26-JUL-1999;	99US-0145698P.
PR	25-JUN-1998;	98US-0090862P.	PR	28-JUL-1999;	99US-0146222P.
PR	26-JUN-1998;	98US-0090863P.	PR	17-AUG-1999;	99US-0149396P.
PR	01-JUL-1998;	98US-0091360P.	PR	15-SEP-1999;	99WO-US021090.
PR	01-JUL-1998;	98US-0091544P.	PR	08-OCT-1999;	99WO-US021547.
PR	02-JUL-1998;	98US-0091478P.	PR	30-NOV-1999;	99WO-US028313.
PR	02-JUL-1998;	98US-0091519P.	PR	01-DEC-1999;	99WO-US028301.
PR	02-JUL-1998;	98US-0091626P.	PR	01-DEC-1999;	99WO-US028634.
PR	02-JUL-1998;	98US-0091628P.	PR	16-DEC-1999;	99WO-US030095.
PR	02-JUL-1998;	98US-0091633P.	PR	20-DEC-1999;	99WO-US030911.
PR	02-JUL-1998;	98US-0091646P.	PR	05-JAN-2000;	2000WO-US000219.
PR	02-JUL-1998;	98US-0091673P.	PR	06-JAN-2000;	2000WO-US000376.
PR	07-JUL-1998;	98US-0091978P.	PR	11-FEB-2000;	2000WO-US003565.
PR	07-JUL-1998;	98US-0091982P.	PR	18-FEB-2000;	2000WO-US004341.
PR	09-JUL-1998;	98US-0092182P.	PR	22-FEB-2000;	2000WO-US004414.
PR	10-JUL-1998;	98US-0092472P.	PR	24-FEB-2000;	2000WO-US004914.
PR	20-JUL-1998;	98US-0093333P.	PR	02-MAR-2000;	2000WO-US005841.
PR	30-JUL-1998;	98US-0094651P.	PR	10-MAR-2000;	2000WO-US006319.
PR	04-AUG-1998;	98US-0095282P.	PR	15-MAR-2000;	2000WO-US006894.
PR	04-AUG-1998;	98US-0095285P.	PR	20-MAR-2000;	2000WO-US007377.
PR	04-AUG-1998;	98US-0095301P.	PR	30-MAR-2000;	2000WO-US008439.
PR	04-AUG-1998;	98US-0095302P.	PR	15-MAY-2000;	2000WO-US013358.
PR	04-AUG-1998;	98US-0095318P.	PR	17-MAY-2000;	2000WO-US013705.
PR	04-AUG-1998;	98US-0095321P.	PR	22-MAY-2000;	2000WO-US014042.
PR	04-AUG-1998;	98US-0095325P.	PR	30-MAY-2000;	2000WO-US014941.
PR	10-AUG-1998;	98US-0095916P.	PR	02-JUN-2000;	2000WO-US015264.
PR	10-AUG-1998;	98US-0095929P.	PR	23-JUN-2000;	2000US-0213637P.
PR	10-AUG-1998;	98US-0096012P.	PR	28-JUL-2000;	2000WO-US020710.
PR	11-AUG-1998;	98US-0096143P.	PR	11-AUG-2000;	2000WO-US022031.
PR	11-AUG-1998;	98US-0096146P.			
PR	12-AUG-1998;	98US-0096329P.			
PR	17-AUG-1998;	98US-0096757P.			
PR	17-AUG-1998;	98US-0096766P.			
PR	17-AUG-1998;	98US-0096768P.			
PR	17-AUG-1998;	98US-0096773P.			
PR	17-AUG-1998;	98US-0096867P.			
PR	17-AUG-1998;	98US-0096891P.			
PR	17-AUG-1998;	98US-0096894P.			
PR	17-AUG-1998;	98US-0096895P.			
PR	17-AUG-1998;	98US-0096897P.			
PR	18-AUG-1998;	98US-0096949P.			
PR	18-AUG-1998;	98US-0096950P.			
PR	18-AUG-1998;	98US-0096959P.			
PR	18-AUG-1998;	98US-0096960P.			
PR	18-AUG-1998;	98US-0097022P.			
PR	19-AUG-1998;	98US-0097141P.			
PR	20-AUG-1998;	98US-0097218P.			
PR	24-AUG-1998;	98US-0097661P.			
PR	26-AUG-1998;	98US-0097952P.			
PR	26-AUG-1998;	98US-0097953P.			
PR	26-AUG-1998;	98US-0097955P.			
PR	26-AUG-1998;	98US-0097971P.			
PR	26-AUG-1998;	98US-0097974P.			
PR	26-AUG-1998;	98US-0097978P.			
PR	26-AUG-1998;	98US-0097979P.			
PR	26-AUG-1998;	98US-0097986P.			
PR	26-AUG-1998;	98US-0098014P.			
PR	31-AUG-1998;	98US-0098525P.			
PR	16-SEP-1998;	98US-0100634P.			
PR	17-SEP-1998;	98WO-US019330.			
PR	17-SEP-1998;	98US-0100858P.			
PR	17-SEP-1998;	98WO-US019437.			
PR	07-OCT-1998;	98WO-US021141.			
PR	01-DEC-1998;	98WO-US025108.			
PR	22-DEC-1998;	98US-0113296P.			
PR	05-JAN-1999;	99WO-US000106.			
PR	08-MAR-1999;	99WO-US005028.			
PR	12-MAR-1999;	99US-0123357P.			
Query Match 100.0%; Score 1031; DB 6; Length 198;			Best Local Similarity 100.0%; Pred. No. 1.4e-104;		
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MATLWGGLRLGSLSLCLALSVLLLAQLSDAAKNFEDVRCKICIPYKENSCHIYNNK 60	QY	1	MATLWGGLRLGSLSLCLALSVLLLAQLSDAAKNFEDVRCKICIPYKENSCHIYNNK 60
Db	1	MATLWGGLRLGSLSLCLALSVLLLAQLSDAAKNFEDVRCKICIPYKENSCHIYNNK 60	Db	1	MATLWGGLRLGSLSLCLALSVLLLAQLSDAAKNFEDVRCKICIPYKENSCHIYNNK 60
QY	61	ISQKDCDCLHVPEPMPVGRGPDVEAYCLRCECKYEERSVTKVTIIYLSILGLLLYV 120	QY	61	ISQKDCDCLHVPEPMPVGRGPDVEAYCLRCECKYEERSVTKVTIIYLSILGLLLYV 120
Db	61	ISQKDCDCLHVPEPMPVGRGPDVEAYCLRCECKYEERSVTKVTIIYLSILGLLLYV 120	Db	61	ISQKDCDCLHVPEPMPVGRGPDVEAYCLRCECKYEERSVTKVTIIYLSILGLLLYV 120
QY	121	YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANAHDVLARSRSANVLNKVEYAOQRWK 180	QY	121	YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANAHDVLARSRSANVLNKVEYAOQRWK 180
Db	121	YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANAHDVLARSRSANVLNKVEYAOQRWK 180	Db	121	YLTVEPILKRLFGHAQLIQSDDDIGDHQPPANAHDVLARSRSANVLNKVEYAOQRWK 180
QY	181	LQVQEQKSVFDRHVLS 198	QY	181	LQVQEQKSVFDRHVLS 198
Db	181	LQVQEQKSVFDRHVLS 198	Db	181	LQVQEQKSVFDRHVLS 198
RESULT 15					
ABU60609					
ID	ASU60609 standard; protein; 198 AA.				
XX	ABU60609;				
XX	01-MAY-2003 (first entry)				
DE	Human secreted/transmembrane protein, #168.				
XX	Human; PRO; secreted; transmembrane; signal peptide; pharmaceutical;				
KW	diagnostic; therapeutic; gene therapy.				
XX	Homo sapiens.				
OS	US2002160384-A1.				
XX	FN				

[illegible][illegible]

Db	61	ISQKDCDLHVBFMPVRGPDVEAYCLRCCKYEERSVTIKVTIIYLSILGLLLLYMV	120
Qy	121	YLTIVEPILKRLFGHAQLIQSDDDIGDHQFFANAHDLVLRSSRANVLNKVEYAQQRWK	180
Db	121	YLTIVEPILKRLFGHAQLIQSDDDIGDHQFFANAHDLVLRSSRANVLNKVEYAQQRWK	180
Qy	181	LQVQQRKSVFDRHVLS	198
Db	181	LQVQQRKSVFDRHVLS	198

Search completed: March 15, 2004, 08:25:47
Job time : 64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 08:20:07 ; Search time 18 Seconds
(without alignments)

572.772 Million cell updates/sec

Title: US-09-997-573-418

Perfect score: 1031

Sequence: 1 MATLWGGLRLGSLLSCL.....WKLVQEQKSVFDRHVLS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1031	100.0	198	1	CK15 HUMAN
2	979.5	95.0	199	1	Q9nq34 homo sapien
3	544	52.8	183	1	CK15 MOUSE
4	541.5	52.5	183	1	TNE9 HUMAN
5	530	51.4	187	1	TNE9 MOUSE
6	226.5	22.0	227	1	PUT2 FUGRU
7	81.5	7.9	606	1	Q9vna4 drosophila
8	80.5	7.8	895	1	ABD4 MOUSE
9	79	7.7	642	1	DSC3 MOUSE
10	78.5	7.6	1700	1	MOQ COREF
11	77.5	7.5	285	1	BAR3 CHITE
12	77.5	7.5	245	1	Y011 MYCPN
13	77.5	7.5	318	1	MOG RAT
14	77.5	7.5	1257	1	NUIM HIPAM
15	77.5	7.5	5703	1	ERB2 RAT
16	76	7.4	318	1	MUSB HUMAN
17	75.5	7.3	315	1	NUIM LEMCA
18	75.5	7.3	318	1	NUIM MOUSE
19	75.5	7.3	402	1	NUIM TAMTE
20	75	7.3	358	1	UT3 YEAST
21	74.5	7.2	318	1	ARGI NEUCR
22	74.5	7.2	1630	1	NUIM PAPHA
23	74	7.2	246	1	ESPI YEAST
24	74	7.2	256	1	MOG BOVIN
25	74	7.2	606	1	TNE9 MOUSE
26	74	7.2	1002	1	ABD4 HUMAN
27	74	7.2	1304	1	SPR AMEPV
28	73.5	7.1	1958	1	CD45 HUMAN
29	73	7.1	318	1	UBR1 SCHPO
30	73	7.1	896	1	NUIM PHACT
31	72.5	7.0	303	1	DSC3 BOVIN
32	72.5	7.0	318	1	Y247 HUMAN
33	72.5	7.0	318	1	NUIM BALMU
					P24967 balaenopter

ALIGNMENTS

RESULT 1

CK15_HUMAN
ID CK15_HUMAN STANDARD; PRT; 198 AA.
AC Q9NQ34;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein Cllorf15.
GN Cllorf15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21418998; PubMed=11528127;
RA Amid C., Bahr A., Mujica A., Sampson N., Bikar S.E., Winterpacht A., Zabel B., Hankeln T., Schmidt E.R.;
RT "Comparative genomic sequencing reveals a strikingly similar architecture of a conserved syntenic region on human chromosome 11p15.3 (including gene ST5) and mouse chromosome 7.";
RT Cytogenet. Cell Genet. 93:284-290(2001).
RN [2]
RP SEQUENCE FROM N.A.
TX TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schusterberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Klausner R.D., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Altschul S.F., Zeeberg B., Moore T., Max S.I., Wang J., Hsieh F., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Diatchenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBCELLULAR LOCATION: Membrane-associated protein (Potential).
CC -!- SIMILARITY: Belongs to the TMEM9 family.

34 72 7.0 584 1 RECN SYNY3
35 72 7.0 1207 1 EGF HUMAN
36 71.5 6.9 318 1 NUIM BRAVA
37 71.5 6.9 323 1 NUIM PARLI
38 71.5 6.9 1264 1 CYAS RABIT
39 71.5 6.9 2482 1 VWF PIG
40 70.5 6.8 1064 1 YY08 METJA
41 70.5 6.8 1194 1 MGRI HUMAN
42 70 6.8 318 1 NUIM MONDO
43 70 6.8 381 1 COS8 YEAST
44 70 6.8 495 1 RA18 HUMAN
45 70 6.8 720 1 YD71 SCHPO

P74374 synechocyst
P01133 homo sapien
O78701 bradyopus va
P12772 paracentrot
P40144 cryotolagus
Q28833 sus scrofa
Q60307 methanococc
O13255 homo sapien
Q78703 monodelphis
P38723 saccharomyc
Q9ns91 homo sapien
Q10326 schizosacch

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).


```
CC -----
DR EMBL; AJ400877; CAB92287.1; -
DR EMBL; BC040124; AAH40124.1; -
DR Genew; HGNC:1168; C11orf15.
DR InterPro; IPR004153; CXCC repeat.
DR InterPro; IPR008853; TMEM9.
DR Pfam; PF03128; CXCC; 1.
DR Pfam; PF05434; TMEM9; 1.
DR Transmembrane.
KW Transmembrane.
FT TRANSMEM 7 29 POTENTIAL.
SQ SEQUENCE 103 125 POTENTIAL.
SQ SEQUENCE 198 AA; 22531 MW; 888E5536A01B44F CRC64;

Query Match 100.08; Score 1031; DB 1; Length 198;
Best Local Similarity 100.08; Pred. No. 4e-94;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLWGGLLRLGSLLSCLALSLLVLAQLSDAANKFEDVRCKICPPYKENSCHIYKN 60
DB 1 MATLWGGLLRLGSLLSCLALSLLVLAQLSDAANKFEDVRCKICPPYKENSCHIYKN 60
QY 61 ISQKDCDCLHVEPMPVPGDVEAYCLRCCKYEERSVTKVITIIYLSILGLLLYV 120
DB 61 ISQKDCDCLHVEPMPVPGDVEAYCLRCCKYEERSVTKVITIIYLSILGLLLYV 120
QY 121 YLTLPVEPIKRLFGHAQLIQSDDDIGDHPFANAHDVLAHSRANVLNKVEYAQQRW 180
DB 121 YLTLPVEPIKRLFGHAQLIQSDDDIGDHPFANAHDVLAHSRANVLNKVEYAQQRW 180
QY 181 LQVQQRKSVFDRHVLS 198
DB 181 LQVQQRKSVFDRHVLS 198

RESULT 2
CK15_MOUSE STANDARD; PRT; 199 AA.
AC Q9JUR8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein C11orf15 homolog.
GN C11ORF15
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21418998; PubMed=11528127;
RA Amid C., Bahr A., Mujica A., Sampson N., Bikar S.E., Winterpacht A.,
Zabel B., Hankeln T., Schmidt E.R.;
RT "Comparative genomic sequencing reveals a strikingly similar
architecture of a conserved syntenic region on human chromosome
11p15.3 (including gene ST5) and mouse chromosome 7.";
RL Cytogenet. Cell Genet. 93:284-290(2001).
[2]
RN SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Pancreas, and Tongue;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
Saito T., Okazaki Y., Gojobori T., Bonc H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pasole G., Quackenbush J.,
Schrömi L.M., Scubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carrinci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
```

```
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
Hayashizaki Y.,
RA Yashnaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RL "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690(2001).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=2338257; PubMed=12477932;
RA Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K.B., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carrinci P., Prange C.,
RA Raha S.S., Lomuellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Mueny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Richardson S., Worley K.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -|- SUBCELLULAR LOCATION: Membrane-associated protein (Potential).
CC -|- SIMILARITY: Belongs to the TMEM9 family.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AJ400878; CAB92295.1; -
DR EMBL; AK009147; BAB26105.1; -
DR EMBL; AK007430; BAB25033.1; -
DR EMBL; BC002208; AAH02208.1; -
DR InterPro; IPR004153; CXCC repeat.
DR InterPro; IPR008853; TMEM9.
DR Pfam; PF03128; CXCC; 1.
DR Pfam; PF05434; TMEM9; 1.
DR Transmembrane.
KW TRANSMEM 12 34 POTENTIAL.
FT TRANSMEM 104 126 POTENTIAL.
SQ SEQUENCE 199 AA; 22607 MW; 476B6FD1BCECA4E5 CRC64;

Query Match 95.08; Score 979.5; DB 1; Length 199;
Best Local Similarity 94.58; Pred. No. 4.6e-89;
Matches 188; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 1 MATLW-GGLLRGLSLLSCLALSLLVLAQLSDAANKFEDVRCKICPPYKENSCHIYKN 59
DB 1 MASLWCGNLLRLGSLLSCLALSLLVLAQLTGAANKFEDVRCKICPPYKENSCHIYKN 60
QY 60 NISQKDCDCLHVEPMPVPGDVEAYCLRCCKYEERSVTKVITIIYLSILGLLLYV 119
DB 61 NISQKDCDCLHVEPMPVPGDVEAYCLRCCKYEERSVTKVITIIYLSILGLLLYV 120
QY 120 VYLTLPVEPIKRLFGHAQLIQSDDDIGDHPFANAHDVLAHSRANVLNKVEYAQQRW 179
DB 121 VYLTLPVEPIKRLFGHSQLQSDDDVGDHPFANAHDVLAHSRANVLNKVEYAQQRW 180
QY 180 KLVQVQQRKSVFDRHVLS 198
DB 181 KLVQVQQRKSVFDRHVLS 199
```

[illegible]

RA Riboldi,Tunnicliffe G.R., Platzer M., Nyakatura G., Elgar G.S.,
RA Brenner S., Rosenthal A.,
RT "Analysis of the genomic loci of Fugu rubripes homologs of the human
RT disease genes UICAM, G6PD and P53";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Membrane-associated protein (Potential).
CC -!- SIMILARITY: Belongs to the TMEM9 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF026198; AAC15584.1; -
DR F8R; T30536; T30536.
DR GO; GO:0005770; C:late endosome; ISS.
DR GO; GO:0005764; C:lysosome; ISS.
DR InterPro; IPR004153; CXCXC_repeat.
DR InterPro; IPR008853; TMEM9.
DR Pfam; PF03128; CXCXC; 1.
DR Pfam; PF05434; TMEM9; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 10 27 POTENTIAL.
FT TRANSMEM 99 121 POTENTIAL.
FT NON TPR 187 187
SQ SEQUENCE 187 AA; 21460 MW; 51DD101FE4666D3B CRC64;

Query Match 51.48; Score 530; DB 1; Length 187;
Best Local Similarity 57.88; Pred. No. 6.3e-45;
Matches 108; Conservative 29; Mismatches 40; Indels 10; Gaps 4;

QY 14 LLSLSIALSVLLLAQLSDA--AKNFDVRCKICPPYKENSCHIYKNIQKDCDLHV 71
DB 8 LMLAALPMTLFLDLNVGVVQAQSFDDVRCKICPPYKENSCHIYKNIQKDCDLHV 67

QY 72 VEPNVPVGPVVEAYCLCECKYERSVTIKVTLIIYVLGLLLYMYLTLVEPIKR 131
DB 68 VDPNVPVGNVDEAYCLCECKYERSVTIRVTLIIYFLSVGALLYMYLFLLDVEFI-R 126

QY 132 RLFGHAQLIQSDDIGHQPFANAHDLVARSRRAN-VLNKVEYAQQRWKLQVQQRKSV 190
DB 127 KPDPALQTLHNEEDSEDIQ-----QMSGDPARGNTVLRVEGAQQRWKLQVQQRKTV 180

QY 191 FDRHVVL 197
DB 181 FDRHKML 187

RESULT 6
ID YB61_DROME STANDARD; PRT; 227 AA.
AC Q9VNA4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein CG1161.
GN CG1161.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidae; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.B., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busem D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Evangelista C.C., Fertaz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.D., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy J., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhan G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
CC -!- SUBCELLULAR LOCATION: Membrane-associated protein (Potential).
CC -!- SIMILARITY: Belongs to the TMEM9 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE003603; AAF52042.1; -
DR FlyBase; FBgn0037313; CG1161.
DR InterPro; IPR008853; TMEM9.
DR Pfam; PF05434; TMEM9; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 7 24 POTENTIAL.
FT TRANSMEM 135 157 POTENTIAL.
SQ SEQUENCE 227 AA; 24653 MW; 7B3F0B42B0A851EE CRC64;

Query Match 22.04; Score 226.5; DB 1; Length 227;
Best Local Similarity 31.64; Pred. No. 5.2e-15;
Matches 50; Conservative 38; Mismatches 63; Indels 7; Gaps 4;

QY 43 KCICPPYKENSCHIYKNIQKDCDLHVPEP-MPVRGPDVEAYCLCECKYERSVTI 101
DB 75 KVIGQPVLAAAPGKSSNSTTECVACAGALLPRLDANGKEL-PICAECKSHVARTTLI 133

QY 102 KVTIIYISILGLLLYMYLTLVEPIKRRL-FGHAQLIQSDDIGHQPFANAHDLV 160
DB 134 KVVIIIVTWISILYIYMLFMLCMLDPLLNKRVKANYQHTNEDDEPTPLPAVNNQELSA 193

QY 161 RSRSRANVLNKVEYAQQRWKLQVQQRKSVFDRHVVL 198
DB 194 ----RANVLNRVGHQDKWKQVRQRRIHYDRHTMLN 227

RESULT 7
ABD4_MOUSE

ID ABD4_MOUSE STANDARD; PRT; 606 AA.
AC O89016;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ATP-binding cassette, sub-family D, member 4 (Peroxisomal membrane
protein 69) (PMP69) (Peroxisomal membrane protein 1-like) (PMP1-L)
(GP7OR).
GN ABCD4 OR PMP1L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98409337; PubMed=9738957;
RA Holzinger A., Muntau A., Mayerhofer P., Kammerer S., Albet S.,
RT "The mouse gene encoding the peroxisomal membrane protein 1-like
protein (PMP1-L): cDNA cloning, genomic organization and comparative
expression studies.";
RL FEBS Lett. 433:179-183 (1998).
CC -1- SUBUNIT: Homodimer or heterodimer (Potential).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Peroxisomal.
CC -1- SIMILARITY: Belongs to the ABC transporter family. ABC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ001166; CAA04570.1; -;
CC MGD; MGI:1349217; Abcd4.
CC InterPro; IPR003593; AAA_Atpase.
CC InterPro; IPR003439; ABC_transporter.
CC Pfam; PF00005; ABC_tran; 1.
CC ProDom; PD000006; ABC_transporter; 1.
CC SMART; SM00382; AAA_1.
CC PROSITE; PS50929; ABC_TM1F; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
CC ATP-binding; Glycoprotein; Transmembrane; Transport; Peroxisome.
CC TRANSMEM 43 63
CC TRANSMEM 76 96
CC TRANSMEM 190 210
CC TRANSMEM 279 299
CC TRANSMEM 314 334
CC NP_BIND 421 428
CC ATP (POTENTIAL).
CC SEQUENCE 606 AA; 68549 MW; 1EF6F044D665A200 CRC64;
Query Match 7.9%; Score 81.5; DB 1; Length 606;
Best Local Similarity 23.8%; Pred. No. 2.7;
Matches 41; Conservative 25; Mismatches 57; Indels 49; Gaps 8;
QY 7 GLRLGLSLGLSLALSVLLLAQSPDAKNFEDVRCKICPPYK-----ENSGHIYNK--- 59
Db 69 GVLGKDLGDFKALTLLAVTLVLNSTLSFDQFTCNLLYVSWRKDLTEHLHLYFRARV 128
QY 60 -----NLSQKDCCLHVPEMPFGPDVEAYCLRCECKYVERSSVTKV-----TIIIV-- 108
Db 129 YITLVLRDDID-----NDQRTISQDVERPC-----RQLSSVTSKLIISFTIYYTY 176
QY 109 -----LSILGILLL-YMVYITLVEPILKRLFLGHAQLIQSDDDIGD 148
Db 177 QCFQSTGWLGPVSIFFGVFGVGTWVNTKLTMGPIV-----TKLVQOEKLEGD 221
RESULT 8
DSC3_MOUSE
ID DSC3_MOUSE STANDARD; PRT; 895 AA.

P55950; O55110; O55122;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Desmocollin 3 precursor.
GN DSC3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=98049324; PubMed=9389456;
RA Chidgey M.A.J., Yue K.K.M., Gould S., Byrne C., Garrod D.R.;
RT "Changing pattern of desmocollin 3 expression accompanies epidermal
RT organisation during skin development.";
RL Dev. Dyn. 210:315-327 (1997).
RN [2]
RP SEQUENCE OF 709-874 FROM N.A., ALTERNATIVE SPLICING, AND TISSUE
RP SPECIFICITY.
RX MEDLINE=98067789; PubMed=9404003;
RA King I.A., Angst B.D., Hunt D.M., Kruger M., Arnmann J., Buxton R.S.;
RT "Hierarchical expression of desmosomal cadherins during stratified
RT epithelial morphogenesis in the mouse.";
RL Differentiation 62:83-96 (1997).
CC -1- FUNCTION: Component of intercellular desmosome junctions. Involved
CC in the interaction of plaque proteins and intermediate filaments
CC mediating cell-cell adhesion. May contribute to epidermal cell
CC positioning (stratification) by mediating differential
CC adhesiveness between cells that express different isoforms.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=3A;
CC IsoId=P55950-1; Sequence=Displayed;
CC Name=3B;
CC IsoId=P55950-2; Sequence=VSP_000665; VSP_000666;
CC -1- TISSUE SPECIFICITY: First expressed at E13.0 in epithelium of
CC whisker pads and external nares, and in most mature vibrissa
CC follicles. 12 hours later, prominently expressed in whiskers and
CC tactile follicles above the eye. At E14.5, also expressed in
CC developing nails and teeth and, at low levels, in ventral and
CC lateral skin. At E15.5, highly expressed in general body epidermis
CC and at E16.5, detected over entire embryo. In the adult, highly
CC expressed in basal layers of stratified cells.
CC -1- DOMAIN: Calcium may be bound by the cadherin-like repeats
CC (Potential).
CC -1- SIMILARITY: Contains 5 cadherin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y11169; CAA72045.1; -;
CC EMBL; AJ000329; CAA03996.1; -;
CC HSSP; P15116; INCJ.
CC MGD; MGI:1194993; Dsc3.
CC InterPro; IPR002126; Cadherin.
CC InterPro; IPR000233; Cadherin_C_term.
CC Pfam; PF00028; cadherin; 5.
CC Pfam; PF01049; Cadherin_C_term; 1.
CC PRINTS; PR00205; CADHERIN.
CC SMART; SM00112; CA; 5.
CC PROSITE; PS00232; CADHERIN_1; 3.
CC PROSITE; PS02068; CADHERIN_2; 5.
CC Cell adhesion; Glycoprotein; Transmembrane; Repeat; Signal;
CC Cytoskeleton; Calcium-binding; Alternative splicing.
CC SIGNAL 1 31
CC PROPEP 32 135
CC POTENTIAL.

FT CHAIN 136 895 DESMOCOLLIN 3.
 FT DOMAIN 136 695 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 696 715 POTENTIAL.
 FT DOMAIN 717 895 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 136 243 CADHERIN 1.
 FT DOMAIN 244 355 CADHERIN 2.
 FT DOMAIN 356 472 CADHERIN 3.
 FT DOMAIN 473 580 CADHERIN 4.
 FT DOMAIN 581 691 CADHERIN 5.
 FT CARBOHYD 166 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 392 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 547 630 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 630 831 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 831 838 KLVHCNQN -> DSIRHTG (in isoform 3B).
 FT VARSPLIC 831 838 /FTid=VSP_000665.
 FT VARSPLIC 839 895 Missing (in isoform 3B).
 FT VARSPLIC 839 895 /FTid=VSP_000666.
 FT CONFLICT 715 715 V -> I (IN REF. 2).
 FT CONFLICT 750 767 WIIHPHCQQLPRLMTN -> NGFTTHANNSSQGFCTGM
 FT CONFLICT 750 767 (IN REF. 2).
 FT CONFLICT 840 840 N -> D (IN REF. 2).
 FT CONFLICT 873 874 EG -> IK (IN REF. 2).
 FT SEQUENCE 895 AA; 100653 MW; D561E67C57232270 CRC64;
 Query Match 7.8%; Score 80.5; DB 1; Length 895;
 Best Local Similarity 25.7%; Pred. No. 5.2;
 Matches 37; Conservative 17; Mismatches 53; Indels 37; Gaps 5;
 QY 29 QLSDAKNFE-----DVRCKICPPYKENGSHYNNKISQKDCDCLHVVPEMPVGRPDVE 83
 DB 628 QVNTAARLSVQKTADVIQVNPVTVDRAGQSATKILRVNLCDCIH----- 674
 QY 84 AYCLRCCKYBERSVTIKVTIIYLSI-IGLLLLYVYTLVEPIKRLRFOHAQLIOS 142
 DB 675 ----PSQCPLRSRA-GITLGKWAYLAILGIALLFSLVLLTVGVVTARK----- 720
 QY 143 DDDIGDHOPFANADVLARSRA 166
 DB 721 ----GKHPEDLAHENLIISNTA 740
 Query Match 7.8%; Score 80.5; DB 1; Length 895;
 Best Local Similarity 25.7%; Pred. No. 5.2;
 Matches 37; Conservative 17; Mismatches 53; Indels 37; Gaps 5;
 QY 29 QLSDAKNFE-----DVRCKICPPYKENGSHYNNKISQKDCDCLHVVPEMPVGRPDVE 83
 DB 628 QVNTAARLSVQKTADVIQVNPVTVDRAGQSATKILRVNLCDCIH----- 674
 QY 84 AYCLRCCKYBERSVTIKVTIIYLSI-IGLLLLYVYTLVEPIKRLRFOHAQLIOS 142
 DB 675 ----PSQCPLRSRA-GITLGKWAYLAILGIALLFSLVLLTVGVVTARK----- 720
 QY 143 DDDIGDHOPFANADVLARSRA 166
 DB 721 ----GKHPEDLAHENLIISNTA 740
 Query Match 7.8%; Score 80.5; DB 1; Length 895;
 Best Local Similarity 25.7%; Pred. No. 5.2;
 Matches 37; Conservative 17; Mismatches 53; Indels 37; Gaps 5;
 QY 29 QLSDAKNFE-----DVRCKICPPYKENGSHYNNKISQKDCDCLHVVPEMPVGRPDVE 83
 DB 628 QVNTAARLSVQKTADVIQVNPVTVDRAGQSATKILRVNLCDCIH----- 674
 QY 84 AYCLRCCKYBERSVTIKVTIIYLSI-IGLLLLYVYTLVEPIKRLRFOHAQLIOS 142
 DB 675 ----PSQCPLRSRA-GITLGKWAYLAILGIALLFSLVLLTVGVVTARK----- 720
 QY 143 DDDIGDHOPFANADVLARSRA 166
 DB 721 ----GKHPEDLAHENLIISNTA 740

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AP005220; BAC18704.1; -
 CC HAMAP; MF 00212; fused; 1.
 CC InterPro; IPR006231; mal_quin_oxido.
 CC TIGRPFAMs; TIGR01320; mal_quin_oxido; 1.
 CC Oxidoreductase; Tricarboxylic acid cycle; Flavoprotein; FAD;
 CC Complete proteome.
 CC DOMAIN 1 142 UNKNOWN.
 CC FT DOMAIN 143 642 MQO.
 CC SEQUENCE 642 AA; 70125 MW; 19E8271DBF472725 CRC64;
 Query Match 7.7%; Score 79; DB 1; Length 642;
 Best Local Similarity 22.5%; Pred. No. 5.1;
 Matches 42; Conservative 26; Mismatches 55; Indels 64; Gaps 11;
 QY 39 DVRCKICPPYKENGSHYNNKISQKDCDCLHVVPEMPVGRPDVE-----AYCLRCCKY 93
 DB 87 ELRTTCTCTPSRSPGQVFT-----TCRSAIH--PCAGPRVDCVQHAGCVRTTRG 136
 QY 94 EER-----SSVTIKVTII-----IVLSILGLLLYVYTLVEP-----ILKRL 133
 DB 137 DKRKLMSDSPKNAQKVTDEADVIVGVGINSAILGMLRL-----EPWSQVIFERL 190
 QY 134 FOHAQLIOSDDIGDHQPFANA---HDVLAERS-----SRA-NVLNVEYAAQOR 178
 DB 191 DGAQCESS-----PWNAGTGHSCALCENVTPEKNGKIDISKAVNINEKTVSRQF 242
 QY 179 WKLVQVE 185
 DB 243 WAHOOVEE 249
 RESULT 10
 BAR3 CHITE
 ID BAR3 CHITE STANDARD; PRT; 1700 AA.
 AC Q03376;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Balbiani ring protein 3 precursor.
 GN BR3.
 OS Chironomus tentans (Widge).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Chironomidae;
 OC Chironomidae; Chironominae; Chironomus.
 CC NCBI_TaxID=7153;
 [1] SEQUENCE FROM N.A.
 RP TISSUE=Salivary gland;
 RC MEDLINE=90172404; PubMed=1689777;
 RA Paulsson G., Lendahl U., Galli J., Ericsson C., Wieslander L.;
 RT "The Balbiani ring 3 gene in Chironomus tentans has a diverged
 RL J. Mol. Biol. 211:331-349(1990).
 CC -!- FUNCTION: Used by the larvae to construct a supramolecular
 CC structure, the larval tube. Balbiani ring protein 3 could play a
 CC role as a transport protein that binds to other proteins
 CC intracellularly and in the gland lumen in order to prevent these
 CC from forming water-insoluble fibers too early.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Salivary gland.
 CC -!- DOMAIN: Has 82 approximate repeats of Cys-x-Cys-x-Cys.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X52263; CAA36506.1; --
DR PIR; S08157; S08167.
DR HSSP; P15358; 18KZ.
DR InterPro; IPR004153; CXKC_repeat.
DR Pfam; PF03128; CXKC; 71.
KW Repeat; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 1700 BALBIANI RING PROTEIN 3.
SQ SEQUENCE 1700 AA; 186145 MW; 34202B28521B0815 CRC64;
Query Match 7.6%; Score 78.5; DB 1; Length 1700;
Best Local Similarity 24.5%; Pred. No. 17;
Matches 23; Conservative 9; Mismatches 39; Indels 23; Gaps 2;
QY 35 KNFEDVRCCKICPPYK-----ENSGHYNNKHSQKDCDLHV 72
DB 898 KFNDFTCSCGSPGKLDCTGNTKWSAETCTCGGDVNRNGNLKFNDFNLCCQCECKNQ 957
QY 73 EPMVRGPDVEAY-CLRCECKVEERSSVTIKVTI 105
DB 958 ENANCKSPRTWNYDTCCKVCNADSDDCVRFQI 991

RESULT 11
Y011_MYCPN STANDARD; PRT; 285 AA.
AC P75098;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein MG011 homolog (D12_orf285).
GN MPN015 OR MP139.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelsreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae";
RL Nucleic Acids Res. 24:4420-4449 (1996).
CC -1- SIMILARITY: BELONGS TO THE ATP-GRASP SUPERFAMILY. RIMK FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AE000016; AAB95787.1; --
DR PIR; S73465; S73465.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 285 AA; 33446 MW; 1D0F653B3EA5A049F CRC64;
Query Match 7.6%; Score 78; DB 1; Length 285;
Best Local Similarity 28.6%; Pred. No. 2.6;
Matches 22; Conservative 12; Mismatches 20; Indels 3;
QY 124 LVEPILKRUFQHAQLIQSDDTDGHPFANADHVLRSP-----SRANVLNKV----- 172
DB 30 LVDELNKKLVGHILL--DDETADHKHKIKVELIINRSRIDFLTKHFLNSFLINPON 87
QY 173 -----EYAQQRKQLQ 182

Db 88 IVLVANDKYETRWLQK 104

RESULT 12

MOG_RAT
ID MOG_RAT STANDARD; PRT; 245 AA.
AC Q63345;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myelin-oligodendrocyte glycoprotein precursor.
GN MOG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=93085763; PubMed=1453482;
RA Gardiner M.V., Amiguet P., Linington C., Matthieu J.-M.;
RT "Myelin/oligodendrocyte glycoprotein is a unique member of the
RT immunoglobulin superfamily";
RL J. Neurosci. Res. 33:177-187 (1992).
RN [2]
RP SEQUENCE OF 28-245 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93376728; PubMed=8367453;
RA Pham-Dinh D., Mattei M.-G., Nussbaum J.-L., Roussel G.,
RA Pontarotti P., Roeckel N., Mather I.H., Artzt K., Lindahl K.F.,
RA Dautigny A.;
RT "Myelin/oligodendrocyte glycoprotein is a member of a subset of the
RT immunoglobulin superfamily encoded within the major
RT histocompatibility complex";
RL Proc. Natl. Acad. Sci. U.S.A. 90:7990-7994 (1993).
RN [3]
RP STRUCTURE BY NMR OF 62-82.
RX MEDLINE=97354172; PubMed=9210466;
RA Albouze-Abou S., Wilson J.C., Bernard C.C.A., von Itzstein M.;
RT "A conformational study of the human and rat encephalitogenic myelin
RT oligodendrocyte glycoprotein peptides 35-55";
RL Eur. J. Biochem. 246:59-70 (1997).
CC -1- FUNCTION: Minor component of the myelin sheath. May be involved in
CC completion and/or maintenance of the myelin sheath and in cell-
CC cell communication.
CC -1- SUBUNIT: May form homodimers.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Found exclusively in the CNS, where it is
CC localized on the surface of myelin and oligodendrocyte cytoplasmic
CC membranes.
CC -1- DEVELOPMENTAL STAGE: A PEAK OF EXPRESSION HAS BEEN OBSERVED
CC BETWEEN POSTNATAL DAYS 15 AND 25, COINCIDENT WITH THE PERIOD OF
CC ACTIVE MYELINATION.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG
CC family.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- CAUTION: Do not confuse myelin-oligodendrocyte glycoprotein (MOG)
CC with oligodendrocyte-myelin glycoprotein (OMG).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M99485; AAA41628.1; --
DR EMBL; L21995; AAF74786.1; --
DR PIR; B47712; B47712.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.

DR PROSITE; P50835; IG_LIKE; 1.
 KW Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 245 MYELIN-OLIGODENDROCYTE GLYCOPROTEIN.
 FT TRANSMEM 156 176 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 177 208 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 209 229 POTENTIAL.
 FT DOMAIN 230 245 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 30 139 IG-LIKE.
 FT DISULFID 51 125 POTENTIAL.
 FT CARBOHYD 58 58 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 245 AA; 27881 MW; C97F8AD6D6A32B4 CRC64;

Query Match 7.5%; Score 77.5; DB 1; Length 245;
 Best Local Similarity 24.9%; Pred. No. 2.5; Indels 77; Gaps 10;
 Matches 49; Conservative 18; Mismatches 53; Indels 77; Gaps 10;

QY 1 MATLWGGLRLGSLLSLCLSLALLAQLSDA-AKNE-----EDVRCKCI 45
 DB 1 MAGWVS-----LSLPSCLLSLLLLQLSRSYAGQFRVIGPHPIRALVGDDELPCR 52
 QY 46 CPPIKENS-----HIYNKVISQKDCCLHVVEMP----- 76
 DB 53 ISPGKNTMEVGMVRSFPRVHLVRNG---KDQD-----AEQAPYRGRTLLKESIGE 105
 QY 77 -----VRGPDVEAY-CLRCCKYEERSVTIKVTIYY-----LSILGL--LLLY 118
 DB 106 GKVALRQTVRFSDGGYTCFFRDHSQEEAAVELKVEDPFYINFGVLLIALVPMVLL 165
 QY 119 MVYTLVEPIKRLRFG 135
 DB 166 QVSVGLVFLQLHRLRG 182

RESULT 13
 NUIIM HIPAM
 ID NUIIM HIPAM STANDARD; PRT; 318 AA.
 AC Q92Z21;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).
 GN MTND1 OR ND1 OR NADH1.
 OS Hippopotamus amphibius (Hippopotamus).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.
 OX NCBI_TaxID=9833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99098146; PubMed=9881471;
 RA Ursing B.M., Arnason U.;
 RT "Analyses of mitochondrial genomes strongly support a hippopotamus-
 whale clade.";
 RL Proc. R. Soc. Lond., B, Biol. Sci. 265:2251-2255(1998).
 CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -!- SIMILARITY: Belongs to the complex I subunit 1 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ010957; CAA09428.1; --
 DR InterPro; IPR001694; Resp_NADH_dh1.
 DR Pfam; PF00146; NADHdh; 1.
 DR PROSITE; PS00667; COMPLEX1_ND1_1; 1.
 DR PROSITE; PS00668; COMPLEX1_ND1_2; 1.
 KW Ubiquinone; Mitochondrion; Oxidoreductase; NAD; Transmembrane.

SQ SEQUENCE 318 AA; 35928 MW; 2B63BBF6E3125B03 CRC64;

Query Match 7.5%; Score 77.5; DB 1; Length 318;
 Best Local Similarity 38.9%; Pred. No. 3.3; Indels 7; Gaps 2;
 Matches 21; Conservative 12; Mismatches 14; Indels 7; Gaps 2;

QY 105 IIVLSILGLLLMYVLTVEPIKRLRFGHAQLIQSDDDDIGH--QPFANA 155
 DB 3 IINTLMVAPILLAMAFITLVE-----RKILGYMQLRKGNVIGPYGLIQPFADA 52

RESULT 14
 ERB2_RAT
 ID ERB2_RAT STANDARD; PRT; 1257 AA.
 AC P06434;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Epidermal growth factor
 receptor-related protein).
 DE ERB2 OR NEU.
 GN Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Neuroblastoma;
 RX MEDLINE=86113662; PubMed=3945311;
 RA Bergmann C.I., Hung M.-C., Weinberg R.A.;
 RT "The neu oncogene encodes an epidermal growth factor receptor-related
 protein.";
 RL Nature 319:226-230(1986).
 RN [2]
 RP SEQUENCE OF 852-905 FROM N.A.
 RX TISSUE=Sciatic nerve;
 RC MEDLINE=91222560; PubMed=2025425;
 RA Lai C., Lemke G.;
 RT "An extended family of protein-tyrosine kinase genes differentially
 expressed in the vertebrate nervous system.";
 RL Neuron 6:691-704(1991).
 RN [3]
 RP STRUCTURE BY NMR OF 650-668.
 RX MEDLINE=92155181; PubMed=1346763;
 RA Gullick W.J., Bottomley A.C., Loftis F.J., Doak D.G., Mulvey D.,
 RA Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;
 RT "Three dimensional structure of the transmembrane region of the proto-
 oncogenic and oncogenic forms of the neu protein.";
 RL EMBO J. 11:43-48(1992).
 CC -!- FUNCTION: Essential component of a neuroregulin-receptor complex,
 although neuroregulins do not interact with it alone. GP30 is a
 potential ligand for this receptor. Not activated by EGF, TGF-
 alpha and amphiregulin.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 tyrosine phosphate.
 CC -!- SUBUNIT: Heterodimer with each of the other ERBB receptors. The
 constitutively activated oncogenic variant forms a homodimer.
 CC Interacts with PRKCA/BP (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- PTM: Ligand-binding increases phosphorylation on tyrosine
 residues (By similarity).
 CC -!- SIMILARITY: Belongs to the EGF receptor family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X03362; CAA27059.1; ALT_INIT.

DR PIR; A24562; TVRTNU.
DR PDB; 1IIJ; 27-JUN-01.
DR PDB; 1N8Y; 18-FEB-03.
DR InterPro; IPR00494; EGFR L domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow Fac recep.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR InterPro; IPR004019; YLP motif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; kinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR Pfam; PF02757; YLP_2.
DR PRINTS; SM00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 4.
DR SMART; SM00219; TyRK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transmembrane; Glycoprotein; MultiGene family; Receptor; Signal;
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW Proto-oncogene; Disease mutation; 3D-structure.
FT SIGNAL 1 21
FT CHAIN 1 21
FT DOMAIN 22 1257
FT TRANSMEM 22 654
FT DOMAIN 655 677
FT DOMAIN 678 1257
FT DOMAIN 139 339
FT DOMAIN 473 646
FT DOMAIN 722 989
FT NP_BIND 728 736
FT BINDING 755 755
FT ACT_SITE 847 847
FT DISULFID 196 205
FT DISULFID 200 213
FT DISULFID 221 228
FT DISULFID 225 236
FT DISULFID 237 245
FT DISULFID 241 253
FT DISULFID 256 265
FT DISULFID 269 296
FT DISULFID 300 312
FT DISULFID 316 332
FT DISULFID 335 339
FT DISULFID 513 522
FT DISULFID 517 530
FT DISULFID 533 542
FT DISULFID 546 562
FT DISULFID 565 578
FT DISULFID 569 586
FT DISULFID 589 598
FT DISULFID 602 625
FT DISULFID 628 636
FT DISULFID 632 644
FT MOD_RES 1141 1141
FT MOD_RES 1250 1250
FT CARBOHYD 68 68
FT CARBOHYD 188 188
FT CARBOHYD 260 260
FT CARBOHYD 532 532
FT CARBOHYD 573 573
FT CARBOHYD 631 631
FT VARIANT 661 661
SQ SEQUENCE 1257 AA; 138931 MW; 61.29264583011402 CRC64;
Query Match 7.5%; Score 77.5; DB 1; Length 1257;
Best Local Similarity 27.4%; Pred. No. 15;
Matches 29; Conservative 19; Mismatches 25; Indels 33; Gaps 7;
QY 46 CPYKNSGHIYKNKISQKDCDLHVPE-----MPV-RGPDEAYCLRCE----- 90

DB 589 CAHYKDSKSCV-----ARCPGKVKPDLSPYWKYPDEGICQPCINCTHSCVDL 639
QY 91 ----CKYERSGVTKVITIIIVLSILGLLLIMVYLIVELIKRR 132
DB 640 DERGCPAEQAS---PVTFI-ATVVG-VLLFLILVVVVVGLIKRR 680
RESULT 15
MUC5B_HUMAN STANDARD; PRT: 5703 AA
ID MUC5B_HUMAN STANDARD; PRT: 5703 AA
AC Q9HC84; O00447; O00573; O14985; O15494; O95291; O95451; Q14881;
AC Q99552; Q9UE28;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High
DE molecular weight salivary mucin M61) (Sublingual gland mucin).
GN MUC5B OR MUC5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 1-1594 FROM N.A.
RA Chen Y.; Di Y.P.; Wu R.;
RT "Molecular cloning of the amino-terminal and 5'-flanking region of the
RT human MUC5B mucin gene.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-1325 FROM N.A.
RA MEDLINE=99009274; PubMed=9790959;
RX Offner G.D., Nunes D.P., Keates A.C., Afzal N.H., Troxler R.F.;
RA "The amino-terminal sequence of MUC5B contains conserved
RT multifunctional D domains: implications for tissue-specific mucin
RT functions.";
RL Biochem. Biophys. Res. Commun. 251:350-355(1998).
RN [3]
RP SEQUENCE OF 40-1324 FROM N.A.
RX MEDLINE=99023932; PubMed=9804771;
RA Desseyn J.-L., Eulsine M.P., Porchet N., Aubert J.-P., Laine A.;
RT "Genomic organization of the human mucin gene MUC5B: cDNA and genomic
RT sequences upstream of the large central exon.";
RL J. Biol. Chem. 273:30157-30164(1998).
RN [4]
RP SEQUENCE OF 1326-4895 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97166151; PubMed=9013550;
RA Desseyn J.-L., Guyonnet-Duperat V., Porchet N., Aubert J.-P.,
RA Laine A.;
RT "Human mucin gene MUC5B, the 10.7 kb large central exon encodes
RT various alternate subdomains resulting in a super-repeat. Structural
RT evidence for a 11p15.5 gene family.";
RL J. Biol. Chem. 272:3168-3178(1997).
RN [5]
RP SEQUENCE OF 4057-4480 FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=97292540; PubMed=9147051;
RA Nielsen P.A., Bennett E.P., Wandall H.H., Therkildsen M.H.,
RA Hannibal J., Clausen H.;
RT "Identification of a major human high molecular weight salivary mucin
RT (M61) as tracheobronchial mucin MUC5B.";
RL Glycobiology 7:413-419(1997).
RN [6]
RP SEQUENCE OF 4721-5703 FROM N.A.
RC TISSUE=Gall bladder;
RX MEDLINE=97293229; PubMed=9164870;
RA Keates A.C., Nunes D.P., Afzal N.H., Troxler R.F., Offner G.D.;
RT "Molecular cloning of a major human gall bladder mucin: complete C-
RT terminal sequence and genomic organization of MUC5B.";
RL Biochem. J. 324:295-303(1997).
RN [7]
RP SEQUENCE OF 4809-5687 FROM N.A.

[illegible]

QY 48 P--YKENSHTYNKNISQDCDCLH---VVEPMFVRGPDVEAYCLRCECK 92
| : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 840 PGLVDSGGGC----IAEEDPCVNNEATYKP-----GETIRVDCNTCTCR 881

Search completed: March 15, 2004, 08:26:15
Job time : 20 secs


```

DDB      191 RHVVL5 196          |||||
RESULT 2
Q7Z649   PRELIMINARY;       PRT;    124 AA.
AAC Q7Z649;
AC 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBurel. 25, Last sequence update)
DD 01-OCT-2003 (TrEMBurel. 25, Last annotation update)
DE C1orf15 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RSC SEQUENCE FROM N.A.
RSP TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RC Strausberg R.;
RD Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RE EMBL; BC015884; AAAH15884.1; -.
RS SEQUENCE 124 AA; 14428 MW; 1D0CC222276D6B62 CRC64;

Query Match        61.6%; Score 635; DB 4; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.2e-61;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 MFVRGPDVEAYCLRCCKCYEERSVTIKVIIILYSILGLLLLYMVYLTLVEPIKRLRF 134
DB 1 MFVRGPDVEAYCLRCCKCYEERSVTIKVIIILYSILGLLLLYMVYLTLVEPIKRLRF 60

QY 135 GHAAQLIQSDDDIGDHOPFANAHVDIARSGRANVLNKVEYAQRWKLQVGEQRKSVPDRH 194
DB 61 GHAAQLIQSDDDIGDHOPFANAHVDIARSGRANVLNKVEYAQRWKLQVGEQRKSVPDRH 120

QY 195 VVLS 198
DB 121 VVLS 124

RESULT 3
Q7ZW08   PRELIMINARY;       PRT;    186 AA.
AC Q7ZW08;
AC 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBurel. 24, Last sequence update)
DD 01-OCT-2003 (TrEMBurel. 25, Last annotation update)
DE Uncharacterized protein.
```



```
DR InterPro; IPR003885; LRR_cyst.
DR InterPro; IPR007092; LRR_SDS22.
DR Pfam; PF00560; LRR; 5.
DR SMART; SMO0365; LRR_SD22; 5.
KW Hypothetical protein.
SQ SEQUENCE 1782 AA; 214024 MW; 8DD60393FDC2CB50 CRC64;

Query Match      8.8%; Score 91; DB 5; Length 1782;
Best Local Similarity 21.0%; Pred. No. 1.2;
Matches 49; Conservative 36; Mismatches 64; Indels 84; Gaps 13;

QY 32 DAAKNFEDVRCKICP-PYKENSCHIYNK-----NISKQDCDCLHV 72
DB 387 DHPREFENMMWK-KCPNDYKNI-EIKNKEQINNSVIKRIEATSKSLNINIGKDDLNI 444
QY 73 EPM--PVRGPPVEAYCLRCEKYEERSV-TIKVTI-----LIY 113
DB 445 EDKEEYTTDVTNCTNNNNCTIYDKKYNKTIYENNIENKIFPFADKLINELMNIVS 504
QY 114 LLLLYMYLTVLVEPIKR-----RLFG-----HAQIQSD 143
DB 505 LNNLIEIFKTNISKIKKINDNITDYFKGLDELEENFTQILNPFSEVKNNDHLSIVLSE 564
QY 144 DDIGHQPFANAHVLAERSRANVLKVE-YAQRW-----KLVQVQRKSVF 191
DB 565 DEINEYSYKN-----YRSNMNLEDYISSYKNKAGSILIEKKKHLF 608

RESULT 8
Q9B9D8 PRELIMINARY; PRT; 119 AA.
AC Q9B9D8;
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE NADH dehydrogenase subunit 1 (EC 1.6.5.3) (NADH-ubiquinone
oxidoreductase chain 1) (Fragment).
GN NDI.
OS Spodoptera littoralis (Egyptian cotton leafworm).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Amphipyrinae; Spodoptera.
OX NCBI_TaxID=7109;
RN [1]
RP SEQUENCE FROM N.A.
RA Abraham D., Ryrholm N., Wittzell H., Scoble M.J., Holloway J.D.,
RA Lofstedt C.;
RT "Molecular phylogeny of the subfamilies in Geometridae (Geometroidea:
RT Lepidoptera)".
RL Mol. Phylogenet. Evol. 0:0-0(2001).
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001694; Resp_NADH_dhl.
DR Pfam; PF00146; NADHdh; 1.
DR PROSITE; PS00667; COMPLEX1_ND1_1; 1.
DR PROSITE; PS00668; COMPLEX1_ND1_2; 1.
KW NAD; Oxidoreductase; Transmembrane; Ubiquinone; Mitochondrion.
FT NON_TER 119
SQ SEQUENCE 119 AA; 13739 MW; 3AC1FA52853252C3 CRC64;

Query Match      8.2%; Score 84.5; DB 8; Length 119;
Best Local Similarity 29.9%; Pred. No. 0.31;
Matches 23; Conservative 18; Mismatches 23; Indels 13; Gaps 3;

QY 107 IYLSILGLLLY-----MVYLTVEPIKRLRFGHAQIQSDDDIGDH---QPFANAH 157
DB 4 IYMFGLLLIILIGLVGVAYLTLE-----RKVLGYIQRKGNKVGFMGILQPFSDAIK 59
QY 158 VLARSRRANVLNKVEY 174
DB 60 LFTKEQTYPNFSNLSY 76

RESULT 10
Q8WRC9 PRELIMINARY; PRT; 1453 AA.
AC Q8WRC9;
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Cysteine repeat modular protein 4 PDCRM4 (Fragment).
OS Plasmodium berghei (strain Anka).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5923;

Query Match      8.2%; Score 84.5; DB 8; Length 235;
Best Local Similarity 29.9%; Pred. No. 0.66;
Matches 23; Conservative 18; Mismatches 23; Indels 13; Gaps 3;

QY 107 IYLSILGLLLY-----MVYLTVEPIKRLRFGHAQIQSDDDIGDH---QPFANAH 157
DB 4 IYMFGLLLIILIGLVGVAYLTLE-----RKVLGYIQRKGNKVGFMGILQPFSDAIK 59
QY 158 VLARSRRANVLNKVEY 174
DB 60 LFTKEQTYPNFSNLSY 76
```

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=ANKA;
RA  Thompson J.;
RT  "Plasmodium berghei Cysteine Repeat Modular Protein 4: PBCRM4
RL  Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF449199; AAL47159.1; -.
DR  InterPro; IPR006209; BGF-like.
DR  PROSITE; PS01186; BGF_2; 2.
FT  NON_TER 1
FT  NON_TER 1453
SQ  SEQUENCE 1453 AA; 169635 MW; 69D66FBB26139AFC CRC64;

Query Match      8.1%; Score 84; DB 5; Length 1453;
Best Local Similarity 20.6%; Pred. No. 5.8;
Matches 44; Conservative 31; Mismatches 61; Indels 78; Gaps 10;

QY  10 RLGLSLSLCLALSVLLLAQSDAANKFEDVRCCKI-----CP-----47
DQ  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  154 RVGNFI---CIKGGKINESIIIGATYDEMCKIEGYMMNOKCILCKDHYCFNGI 210
QY  48 -----PYKEN-----SGHYKNKISQKDCCLHVVPEPVRGPDVEAYCLRC 89
DQ  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  211 ITQCKNETSQIGSKLCKENCVCCKENFIYK-----ENNCTYCVSIKVKSEPI--YCKLC 264
QY  90 ECKYERSVTIKVTIYIISLGLLL---YMYLTLPILKRLFLG-----HAQLI 140
DQ  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  265 EEKY-----LNINIFVPHQNNYGLVNTIHHYVEDFTLSRENVYAYFGCKLERENREI 319
QY  141 QSDDDIGHQPPANAHVLAARSRRANVKNVEY 174
DQ  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  320 QIDSKI-----PRNYANYINPIEF 338

RESULT 11
Q9B9D2 PRELIMINARY; PRT; 115 AA.
AC Q9B9D2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE NADH dehydrogenase subunit 1 (EC 1.6.5.3) (NADH-ubiquinone
oxidoreductase chain 1) (Fragment).
GN NDI.
OS Scapula ornata.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Geometroidea; Geometridae; Sterrhinae; Scopula.
OX NCBI_TaxID=104445;
RN [1]
RP SEQUENCE FROM N.A.
RA Abraham D., Ryrholm N., Wittzell H., Scoble M.J., Holloway J.D.,
RA Lotstedt C.;
RT "Molecular phylogeny of the subfamilies in Geometridae (Geometroidea:
Lepidoptera).";
RL Mol. Phylogenet. Evol. 0:0-0(2001).
CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
DR EMBL; AF178863; AAK00990.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001694; Resp_NADH_dhl.
DR Pfam; PF00146; NADHdh; 1.
DR PROSITE; PS00667; COMPLEX1_NDI_1; 1.
KW NAD; Oxidoreductase; Transmembrane; Ubiquinone; Mitochondrion.
FT NON_TER 1
FT NON_TER 115
SQ SEQUENCE 115 AA; 13285 MW; AB478F6F41F855EE CRC64;

Query Match      8.1%; Score 83.5; DB 8; Length 318;
Best Local Similarity 41.1%; Pred. No. 1.2;
Matches 23; Conservative 9; Mismatches 11; Indels 13; Gaps 3;

QY  109 LSLTGLLLY-----MYLTLPILKRLFGHAQLIQSDDDIGDH---QPFANA 155
DQ  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  1 MTMNLILLIMTSLAMAFLLTVE-----KLLGHQKLRKGNVGVPHGLQPFADA 52

RESULT 13
Q7XMP2 PRELIMINARY; PRT; 783 AA.
AC Q7XMP2;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
SQ SEQUENCE 318 AA; 35767 MW; E09515DEBB038890 CRC64;

Query Match      8.1%; Score 83.5; DB 8; Length 318;
Best Local Similarity 41.1%; Pred. No. 1.2;
Matches 23; Conservative 9; Mismatches 11; Indels 13; Gaps 3;

QY  109 LSLTGLLLY-----MYLTLPILKRLFGHAQLIQSDDDIGDH---QPFANA 155
DQ  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  1 MTMNLILLIMTSLAMAFLLTVE-----KLLGHQKLRKGNVGVPHGLQPFADA 52
```



```
DE OSJNB0059K02.19 protein.
GN OSJNB0059K02.19.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=4530;
[1]
RN SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L., Li T.,
RA Lu Y.Q., Mu S.L., Liu X.H., Lu T.F., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Yang K., Zhou B., Chen Z.H.,
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu P.H., Hong G.F.,
RA Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL606692; CAB04509.1;
SQ SEQUENCE 783 AA; 86458 MW; 81BCF310A35088F1 CRC64;

Query Match 8.1%; Score 83.5; DB 10; Length 783;
Best Local Similarity 20.8%; Pred. No. 3.3;
Matches 47; Conservative 28; Mismatches 72; Indels 79; Gaps 10;

QY 5 WGLRLGLSLSLCLALSVLLAQLSDAANKFEDVRCKICPPYKNSGHIYNNK-NISQ 63
DB 287 WLDLRDGAIFLS-----VGPNITDFRCLSL-----HSSYYINDLNDR 326
QY 64 KDCDCLHVPEP-----VRGPDV---EAYCLRCECKYBERSSVTIK-----VT 104
DB 327 RCNCSQYEGNPIYRDGCRDDEQCPDVYFCHGTCINMPGTVCRLAKSVKSLPGLIT 386
QY 105 ILYLSILGLLLYMYLVTLVPEI-----LKRRLP-GHA-----QLIQDDDIGDHQP 151
DB 387 IIAVSAGFGLIFSLGVAITNKIKQRRAKLRKFFKXNHLGLLQQLISSNKDIAERMK 446
QY 152 FANAHDVLAERSRANVLNKVEAQQRWKIQVQEQKSVFDRHVL 197
DB 447 I-----FSLELDQATKFDQNRIL 466

RESULT 14
Q869E1 PRELIMINARY; PRT; 1192 AA.
AC Q869E1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to Xenopus laevis (African clawed frog) DNA ligase I (EC
DE 6.5.1.1) (Polydeoxyribonucleotide synthase).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OC NCBI_TaxID=44689;
[1]
RN SEQUENCE FROM N.A.
RA STRAIN=AX4;
RA Gloeckner G., Bichinger L., Szafarski K., Pachebat J., Dear P.,
RA Lehman R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RA "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RA Nature 418:79-85 (2002).
[2]
RN SEQUENCE FROM N.A.
RA STRAIN=AX4;
RA Baumgart C.;
RA Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC116957; AA052526.1;
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0003910; F:DNA ligase (ATP) activity; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
```

3
2
2
4
.
5
4
4
7
7
7
5
5
2
3

Query Match 8.0%; Score 82.5; DB 8; Length 120;

Search completed: March 15, 2004, 08:27:14

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 08:23:48 ; Search time 21 Seconds
(without alignments)
906.949 Million cell updates/sec

Title: US-09-997-573-418
Perfect score: 1031
Sequence: 1 MATLWGLRLGSLSLSLCL.....WKLQVQQRKSVFDRHVLS 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	530	51.4	187	T30536	hypothetical prote
2	277.5	26.9	261	T16732	hypothetical prote
3	84.5	8.2	235	A4233	NADH2 dehydrogenas
4	80.5	7.8	432	T37684	probable 6-phospho
5	79	7.7	479	D86340	protein F2D10.30 [
6	78.5	7.6	450	AF0207	probable membrane
7	78.5	7.6	841	T01011	hypothetical prote
8	78.5	7.6	1700	S08167	Balbani ring 3 pr
9	78	7.6	285	S73465	MG011 homolog D12
10	77.5	7.5	275	S75818	hypothetical prote
11	76.5	7.4	2471	T42977	large tegument pro
12	75.5	7.3	315	QXMSIM	NADH2 dehydrogenas
13	75.5	7.3	401	A3537	oligopeptide trans
14	75.5	7.3	402	S46641	probable membrane
15	75.5	7.3	617	S48160	metalloproteinase
16	75	7.3	297	T08589	hypothetical prote
17	75	7.3	325	T47229	arginase (EC 3.5.3
18	75	7.3	358	T47228	NADH2 dehydrogenas
19	74.5	7.2	318	T11506	hypothetical prote
20	74.5	7.2	1372	T25933	ESPI protein - yea
21	74.5	7.2	1630	S64403	myelin/oligodendro
22	74	7.2	246	A47712	T-cell antigen 4-1
23	74	7.2	256	B32393	ABC-transporting p
24	74	7.2	606	JC5604	spheroidin precurs
25	74	7.2	1003	1 PYVZAM	leukocyte common a
26	74	7.2	1304	1 A46546	protein-tyrosine k
27	73.5	7.1	1260	1 TVRTNU	hypothetical prote
28	73.5	7.1	1958	1 T39808	hypothetical prote
29	73	7.1	527	2 T23699	hypothetical prote

30 73 7.1 599 2 T25835 hypothetical prote
31 73 7.1 896 2 I45858 desmocollin - bovi
32 72.5 7.0 318 2 A58850 NADH2 dehydrogenas
33 72.5 7.0 318 2 S41820 probable transcript
34 72.5 7.0 350 2 C83608 hypothetical prote
35 72.5 7.0 351 2 AC2289 probable receptor-
36 72.5 7.0 735 2 T00850 hypothetical prote
37 72 7.0 584 2 S76210 epidermal growth f
38 72 7.0 1207 1 EGHU hypothetical prote
39 72 7.0 1582 2 T15308 proline-rich pepti
40 72 7.0 5762 2 A41819 NADH2 dehydrogenas
41 71.5 6.9 323 2 A34284 type V adenylyl cy
42 71.5 6.9 1264 2 S41603 mucin MUC5B, trach
43 71.5 6.9 1321 2 JE0352 two-component sens
44 71 6.9 481 2 AE1207 hypothetical prote
45 71 6.9 940 2 T31575 hypothetical prote

ALIGNMENTS

RESULT 1

T30536 hypothetical protein 2 - Fugu rubripes (fragment)

C:Species: Fugu rubripes

C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000

C:Accession: T30536

R:Riboldi Tunncliffe, G.R.; Platzer, M.; Nyakatura, G.; Elgar, G.S.; Brenner, S.; Ros
submitted to the EMBL Data Library, September 1997

A:Description: Analysis of the genomic loci of Fugu rubripes homologs of the human disc

A:Reference number: Z20848

A:Accession: T30536

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-187 <RIB>

A:Cross-references: EMBL:AF026198; NID:g3098263; PID:g3098268; PID:AA015584.1

C:Genetics:

A>Note: PUT2

A>Note: Intron positions not resolved (incomplete sequence)

Query Match 51.4%; Score 530; DB 2; Length 187;

Best Local Similarity 57.8%; Pred. No. 3.7e-43;

Matches 108; Conservative 29; Mismatches 40; Indels 10; Gaps 4;

QY 14 LLSLSCLSLVLLLAQLSDA--AKNFEDVRCKICPPYKNSGHLYNKNISQKDCDLHV 71

Db 8 LWMLEAFALMTLFLDNGVGTQAKSFDDVRCKICPPYRNISGHLYNKNFTQKDCNCLHV 67

QY 72 VBPMPVRGPDVEAYCLRCCKEYERSSVTIKVIIISILGLLLLMVYLTLPVPIKR 131

Db 68 VDPMPVPGNDVEAYCLLCKEYERSTWIRVIIIFLSVVGALLLYMLFLLDPLI-R 126

QY 132 RLFGHAQLIQSDDDIGDHQPPANAHDLVARSRRAN-VLNKVEYAOQKQVQQRKSV 190

Db 127 KPDPLAQLTHNEEDSEDIQP-----QMSGDPARGNTVLERVEGAQQRWKQVQQRKTIV 180

QY 191 FDRHVVL 197

Db 181 FDRHML 187

RESULT 2

T16732 hypothetical protein R12C12.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999

C:Accession: T16732

R:Favella, T.

submitted to the EMBL Data Library, July 1995

A:Description: The sequence of C. elegans cosmid R12C12.

A:Reference number: Z18568

A:Accession: T16732

A>Status: preliminary; translated from GB/EMBL/DBJ

submitted to the EMBL Data Library, October 1999
A:Reference number: Z21736
A:Accession: T37684
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-432 <MCD>
A:Cross-references: EMBL:ALJ32675; PIDN:CAB59697.1; GSPDB:GN00066; SPDB:SPAC144.17C
A:Introns: 68/3; 115/2; 230/1
A:Superfamily: Caenorhabditis elegans hypothetical protein R12C12.6
C:Superfamily: Caenorhabditis elegans hypothetical protein R12C12.6

Query Match 26.9%; Score 27.5; DB 2; Length 261;
Best Local Similarity 32.1%; Pred. No. 6.2e-19;
Matches 68; Conservative 44; Mismatches 55; Indels 35; Gaps 8;

QY 20 LALSIVLLAQLSDAA--KNFEDVRCKICP-----PYKENS-----GHYINK-NIS 62
DB 52 LFIALFLPALSQAGTEANFEDTRCICPCLLKFLDLAENTTEKTEGLRRRFYTKNIE 111
QY 63 QKDCDLHVVEPMPVRGPD---VAYCLRCCKEYESSVTIKVTIILYLSILGLLLYM 119
DB 112 PSHKPSNIVKQVSNFVDETHMDAFLANCRIYESNTVLKVVIFVICVIAVLGYM 171
QY 120 VYLTVLVEPIL-KRRFLGHAQLIQSDDDIGDH-----QPFANAHVDVLARSRA 166
DB 172 VFLMCLDPMLRKRLISVQ--QHNDMEDNIFAAAPSTDDSESSASNSMDTQGTTRS 229
QY 167 NVLNVKEVAAQQRWKLQVQQRKSVFDRHVLS 198
DB 230 NVLGRVEAFQNRWKKVBEQRNIPEDETMLN 261

RESULT 3
A44233
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 1 - fall armyworm mitochondrion (fra
N:Alternate names: NADH-ubiquinone oxidoreductase chain 1
C:Species: mitochondrion Spodoptera frugiperda (fall armyworm)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C:Accession: A44233
R:Pasley, D.P.; Ke, L.D.
Mol. Biol. Evol. 9, 1061-1075, 1992
A:Title: Sequence evolution in mitochondrial ribosomal and ND-1 genes in lepidoptera: im
A:Reference number: A44233; MUID:93061985; PMID:1435234
A:Accession: A44233
A:Molecule type: DNA
A:Residues: 1-235 <PAS>
A:Cross-references: GB:M76713; NID:G343352; PIDN:AAA32079.1; PID:G552886
A:Note: sequence extracted from NCBI backbone (NCBIP:118938)
C:Genetics:
A:Gene: ND-1
A:Genome: mitochondrion
A:Start codon: ATA
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 1
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 8.2%; Score 84.5; DB 2; Length 235;
Best Local Similarity 29.9%; Pred. No. 1.4;
Matches 23; Conservative 18; Mismatches 23; Indels 13; Gaps 3;

QY 107 IYLSILGLLLY-----MVYFLVFPILKRLFGHAQLIQSDDDIGDH---QPFANAH 157
DB 4 IYMFGLLLIIGILGVAYLTLE---RKVLGVQIRKGNKVGFMGILQPFSDA 59

QY 158 VLARSRRANVLNKVEY 174
DB 60 LFTKEQTPNFSNLYS 76

RESULT 4
T37684
Probable 6-phosphofructo-2-kinase/fructose-2, 6-bisphosphatase - fission yeast (Schizosa
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C:Accession: T37684
R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Davis, P.; Churcher, C.M.

submitted to the EMBL Data Library, October 1999
A:Reference number: Z21736
A:Accession: T37684
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-432 <MCD>
A:Cross-references: EMBL:ALJ32675; PIDN:CAB59697.1; GSPDB:GN00066; SPDB:SPAC144.17C
A:Introns: 68/3; 115/2; 230/1
A:Superfamily: 6-phosphofructo-2-kinase / fructose-2,6-bisphosphate 2-phosphatase; phos
F:259-430/Domain: phosphoglycerate mutase homology <PGMH>

Query Match 7.8%; Score 80.5; DB 2; Length 432;
Best Local Similarity 25.6%; Pred. No. 6.3;
Matches 45; Conservative 26; Mismatches 50; Indels 55; Gaps 11;

QY 19 CLA-LSVLLAQ-----LSDAANKN--FEDVRCKICPPYKENS 56
DB 123 CLADLEIFLLKEGQVAYDATNGTRRRILYDRFKCGFKILFIESLC-----NKEDV 177
QY 57 YKNNISQKDCCLHVVEPMPVRGPDVE---AYCLRCCKEYESSVTIKV---ERSVTIKV--- 103
DB 178 INANIQ-----EAIHVSEEF--RNWOLEMAEKYCRIDILKCHYETIDEKDYFVKMINF 231
QY 104 --TIIYLSILGLLLYLVVFLVPEILKRRFL-----GHAQLIQSDDDIGDHQ 150
DB 232 AETIANKSNEGYSLLSRILFLMNTLAKRFLVPKASMRPLKREPEDIENRQ 287

RESULT 5
D86340
Protein F2D10.30 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D86340
R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
anssen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.B.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86340
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-479 <STO>
A:Cross-references: GB:AE005172; NID:98886943; PIDN:AAF80629.1; GSPDB:GN00141
C:Genetics:
A:Gene: F2D10.30
A:Map position: 1

Query Match 7.7%; Score 79; DB 2; Length 479;
Best Local Similarity 25.5%; Pred. No. 9.7;
Matches 35; Conservative 17; Mismatches 47; Indels 38; Gaps 6;

QY 14 LLSLSCLALSIVLLAQLSDAANKNFEDVRCKICPPYKEN-----SCHYNNKINSQKDCD 67
DB 13 LLKLDPKSLAMKCTRESINSHISED-----PYKSNLYSLVGFGLLHNSYGSKSLF 65
QY 68 CLHVVEPMPVR---GPDVEAYCLRCCKEYESSVTIKVTIILYLSILGLLLYVYLT 124
DB 66 CNPFGDSMPFRYTVSLDIKTRFL-CSCS-----GULLLFMDYLCV 104

QY 125 VEPILKR-RLFGHAQLI 140
DB 105 ANPLTKRYFLDHSKSI 121

RESULT 6
AF0207
Probable membrane protein YP01702 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AF0207
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11596360
A:Accession: AF0207
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-450 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC90522.1; PID:gl5975733; GSPDB:GN00175
C:Genetics:
C:Superfamily: hypothetical protein H1671
Query Match 7.6%; Score 78.5; DB 2; Length 450;
Best Local Similarity 22.5%; Pred. No. 10;
Matches 23; Conservative 20; Mismatches 46; Indels 13; Gaps 4;
QY 38 EDVRCKICPPYKNSGHIYKNIS-----QKDCCLHVEPMPVPGDVEAYCLRCCKY 93
DB 25 DDILCRATLYPMKI---HAIQRLSTARAQRCCEC-DALFTPLPLGNOTAYCPNCAGI 80
QY 94 EERSSVITKVTIYILSILGLLLYMYLTLVEPILKRLFG 135
DB 81 TSGRDWS-----LTRLTAMACMLLLIPFAFTPLISIRLLG 117
RESULT 7
T01011
Hypothetical protein At2g39810 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein T517.11
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 23-Mar-2001
C:Accession: T01011; G84821
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: G84821
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-841 <STO>
A:Cross-references: GB:AE002093; NID:g2642163; PIDN:AB87130.1; GSPDB:GN00139
C:Genetics:
A:Gene: T517.11; At2g39810
A:Map position: 2
A:Introns: 133/2; 200/3; 365/3; 529/3; 575/3; 788/2
C:Superfamily: Arabidopsis thaliana hypothetical protein At2g39810
Query Match 7.6%; Score 78.5; DB 2; Length 841;
Best Local Similarity 22.7%; Pred. No. 19;
Matches 51; Conservative 27; Mismatches 74; Indels 73; Gaps 11;

QY 1 MATLWGGLRLGLSL--SCLALSILLQAQLDAKKNFEDVRCKICPPYKNSGHIY 58
DB 398 MVLWSGRDGVSELVSGEAVTALRVRVGCLLSEAFYQ-----RTLCLKVKN--L 449
QY 59 KNTSQK-----DCD-----CLHVVPEMPVR-----GPDVEAYCLRC--E 90
DB 450 KNGAVKXASDDLDIWSWTEWELVNEFCCLSRRLVDRRIELPWNPEDEKYLHRCLLD 509
QY 91 CKVEERSVITKVTIYILSILGLLLYMYLTLVEPILKRLRFGHAQLIQSDDDIGDHQ 150
DB 510 SATDPSSAVGSLVLYFLQRYIQAYQVDL-----RLQKIEAFVSDNQIGE-- 558
QY 151 PFANADVTLARSANVLNKVEYAQQRWKLQVQEQKSVFDRHV 195
DB 559 -----EVMFMRSS-----QSHW-----RKELVDRAI 579
RESULT 8
S08167
Balbiani ring 3 protein - midge (Chironomus tentans)
C:Species: Chironomus tentans
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000
C:Accession: S08167
R:Paulsson, G.; Lendahl, U.; Galli, J.; Ericsson, C.; Wieslander, L. J. Mol. Biol. 211, 331-349, 1990
A:Title: The balbiani ring 3 gene in Chironomus tentans has a diverged repetitive stru
A:Reference number: S08167; MUID:90172404; PMID:1689777
A:Accession: S08167
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-1700 <PAU>
A:Cross-references: GB:X52263; NID:g7057; PIDN:CAA36506.1; PID:g7058
C:Genetics:
A:Gene: BR3
A:Map position: 4
C:Superfamily: unassigned Balbiani ring proteins
Query Match 7.6%; Score 78.5; DB 2; Length 1700;
Best Local Similarity 24.5%; Pred. No. 38;
Matches 23; Conservative 9; Mismatches 39; Indels 23; Gaps 2;
QY 35 KNFEDVRCKICPPYK-----ENSGHIYKNISQKDCCLHVV 72
DB 898 KKFNDFTSCGCPGSKLDTGNTKWSAETCTCGGVNRCGNLKNFNDNLQCCECKNQ 957
QY 73 EPMVPGDVEAY-CLRCCKYBERSVITKVTI 105
DB 958 EMANCKSPRTWNYDTCKVCKNADSDDCVKPQI 991
RESULT 9
S73465
MG011 homolog D12_orf285 - Mycoplasma pneumoniae (strain ATCC 29342)
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 10-Dec-1999
C:Accession: S73465
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirk, E.; Li, B.C.; Herrmann, R. Nucleic Acids Res. 24, 4420-4449, 1996
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumonia.
A:Reference number: S73327; MUID:97105885; PMID:8948633
A:Accession: S73465
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-285 <HIM>
A:Cross-references: EMBL:AE000016; GB:U00089; NID:gl673796; PIDN:AAB95787.1; PID:gl673
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: ribosomal protein S6 modification protein rimK
Query Match 7.6%; Score 78; DB 1; Length 285;

Best Local Similarity 28.6%; Pred. No. 7.2;
Matches 22; Conservative 12; Mismatches 23; Indels 20; Gaps 3;

QY 124 LVEPLKRLFGHAQLIQSDDDIGDHPFANAHDVLAARS-----SRANVLNKV----- 172
DB 30 LVDELNKKLVGHILL--DDETADHKHKVVELLINSRRIDPLTKHNFINSFLINPQN 87

QY 173 -----EYAOQRWKLQ 182
DB 88 IVLVANDKYETRYMLKQ 104

RESULT 10
S75818
hypoetical protein slr1276 - *Synechocystis* sp. (strain PCC 6803)
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S75818
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75818
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-275 <KAN>
A:Cross-references: EMBL:D90913; GB:AB001339; NID:G1653348; PIDN:BAAL8277.1; PID:G165336
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: *Synechocystis* hypothetical protein slr1276

Query Match 7.5%; Score 77.5; DB 2; Length 275;
Best Local Similarity 29.7%; Pred. No. 7.8;
Matches 38; Conservative 14; Mismatches 41; Indels 35; Gaps 5;

QY 94 ERSSTVTKVII-----YLSILGILLLYMYVTLVLEPIL 129
DB 6 EYSSST--TIIDEEFDAGNYPTAFGITFTFQVSGISLGLGLANGYLLNFWMPAL 62

QY 130 KRLFGHAQLIQSDDDIGDHPFANAHDVLAARSRAVLAQRWKLQVQERKS 189
DB 63 -----GEYQLQRODEAKQAQVDQSPVLMRLANAQV--QLQEAQR-KATVLELYAN 114

QY 190 VFDRHVVL 197
DB 115 SEDLNTIL 122

RESULT 11
T42977
large tegument protein - ateline herpesvirus 3 (strain 73)
C:Species: ateline herpesvirus 3
A:Variety: strain 73
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C:Accession: T42977
R:Albrecht, J.C.; Fleckenstein, B.
submitted to the EMBL Data Library, August 1998
A:Description: Primary structure of the herpesvirus ateles genome.
A:Reference number: 222274
A:Accession: T42977
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2471 <ALB>
A:Cross-references: EMBL:AF083424; PIDN:AAC95588.1
A:Experimental source: strain 73

Query Match 7.4%; Score 76.5; DB 2; Length 2471;
Best Local Similarity 22.4%; Pred. No. 85;
Matches 38; Conservative 28; Mismatches 81; Indels 23; Gaps 6;

QY 27 LAQLSDAAKNFEDVRC---KCIPEPKYKNSGHYNNKNSKDCDCLHVEPMPVRCPDVE 83

DB 1731 LKASLQKSFQEAATLLHKTLKPLFKVSELTLENNHNSKDCSDALLPFPKWKPID 1790
QY 84 AYCLRE-----CKEERSSTVTKVIIYLSILGILLLYMYVTL-----VEPIKLR 131
DB 1791 IQSVNWEHPSFMSICKNQARARITF---LTLALKIIDPTILNQLWSSLNPAANTSEPTSYS 1847

QY 132 RLFGHAQLIQSDDDIGDHPF---ANAHDLARSRAVLAQRWKLQVYAOQR 178
DB 1848 LUFNY--LIASEFDKTVSTFLEPCNPVPVAYGIQTGNTGNTSYIQOK 1895

RESULT 12

QXMS1M
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 1 - mouse mitochondrion
N:Alternate names: NADH-ubiquinone oxidoreductase chain 1
C:Species: Mitochondrion Mus musculus (house mouse)
C:Date: 02-Apr-1982 #sequence_revision 17-Jul-1998 #text_change 03-Jun-2002
C:Accession: A00409
R:Bibb, M.J.; Van Eeten, R.A.; Wright, C.T.; Walberg, M.W.; Clayton, D.A.
Cell 26, 167-180, 1981
A:Title: Sequence and gene organization of mouse mitochondrial DNA.
A:Reference number: A00153; MUID:82137051; PMID:7332926
A:Accession: A00409
A:Molecule type: DNA
A:Residues: 1-315 <BIB>
A:Cross-references: GB:J01420; NID:G342520; PIDN:AA848644.1; PID:G896295
A:Note: the authors translated the initiation codon ATT for residue 1 as Ile

C:Genetics:
A:Gene: NDI
A:Genome: mitochondrion
A:Genetic code: SGC1
A:Start codon: ATT
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 1
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 7.3%; Score 75.5; DB 1; Length 315;
Best Local Similarity 33.3%; Pred. No. 14;
Matches 20; Conservative 15; Mismatches 18; Indels 7; Gaps 2;

QY 109 LSLGILLLYMYVTLVLEPILKRLFGHAQLIQSDDDIGDCH---QPFANAHDVLAARSR 165
DB 4 LTLVLPILIAAFLTIVE----RKILGYMQLARKPNIVGPILOPFADAKLFMKEPMR 59

RESULT 13

AE3537
oligopeptide transport ATP-binding protein oppF BMEII0223 [imported] - *Brucella meliten*
C:Species: *Brucella melitensis*
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AE3537
R:DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes,
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella meliten*
A:Reference number: AD3252; PMID:11756688
A:Accession: AE3537
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-401 <KUR>
A:Cross-references: GB:AE008918; PIDN:AAL53464.1; PID:G17984365; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEII0223
A:Map position: II

Query Match 7.3%; Score 75.5; DB 2; Length 401;
Best Local Similarity 22.6%; Pred. No. 18;
Matches 48; Conservative 21; Mismatches 76; Indels 67; Gaps 7;

QY 35 KNFEDVRKCIIPPYK-----ENSGHIYNNKISKDCDCLHVEPM----- 75
DB 13 KDALSIRAAQNRFPYAGRKQPIGGSGSIYRRATSPKACRPSNVT-PMQNLWRGALMREN 71

Qy 76 -----PVRG-----PDVEAYCLRCCKYBERSVTIKVIIYLSILGLL 115
Db 72 STBIALSVGRKLSFPVKGTTFIQRDIAAVKAVDDISFDLRKGETLG---LVGESCGKGS 128
Qy 116 LLYWVYLTVEPIIKRLFP-----GHALIQSDDDDIGHOPFANADV 158
Db 129 TLGRCILRLPEPSDQGVLFNGNDMTKLNARDMAARKHLQVFQD-----PFASLHPR 181
Qy 159 LARSRSRANVLNKVEYAQQRWKQLQVQRKSV 190
Db 182 MRINKESIAEPLRISDLTQARKERVQBMRLV 213

RESULT 14

S46641

Probable membrane protein YJL193w - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: hypothetical protein J0349

C;Species: *Saccharomyces cerevisiae*

C;Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 19-Apr-2002

A;Accession: S46641; S56980; S56976

R;Purnelle, B.; Coster, F.; Goffeau, A.

Yeast 10, 1235-1249, 1994

A;Title: The sequence of a 36 kb segment on the left arm of yeast chromosome X identifies

ase gene ACO1 and two homologues to chromosome III genes.

A;Reference number: S46621; MUID:95274326; PMID:7754713

A;Accession: S46641

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-402 <PUR>

A;Cross-references: EMBL:X77688; NID:g1183992; PIDN:CAAS4767.1; PID:g547602

R;Purnelle, B.; Coster, F.; Goffeau, A.

submitted to the Protein Sequence Database, September 1995

A;Reference number: S56977

A;Accession: S56980

A;Molecule type: DNA

A;Residues: 1-402 <PUW>

A;Cross-references: EMBL:Z49468; NID:g1008405; PID:g1008406; MIPS:YJL193w

R;Obermaier, B.; Piravandi, E.; Rinke, M.; Domdey, H.

submitted to the Protein Sequence Database, September 1995

A;Reference number: S56937

A;Accession: S56976

A;Molecule type: DNA

A;Residues: 120-402 <OBE>

A;Cross-references: EMBL:Z49468; MIPS:YJL193w

C;Genetics:

A;Cross-references: SGD:S0003729

A;Map position: 10L

C;Keywords: transmembrane protein

Query Match 7.3%; Score 75.5; DB 2; Length 402;
Best Local Similarity 21.2%; Pred. No.18;
Matches 34; Conservative 22; Mismatches 47; Indels 57; Gaps 5;

Qy 9 LRLGSLLSLSIALSVLLLAQLSDAAKNFEDVRCKICPPYKENSCH----- 55
Db 166 ITLTLFSLCTLVLGWIIQ-----EDNRGPASSNELREFSKYGVICAMISMEIFV 217
Qy 56 ---INKNI-----SOKDCCLHVVEPVRGPDVEAYCLRCCKYERSSVTI-- 101
Db 218 LQNIYKGVFTVRSQTSQNSGFSRQESPLPL-----YEKLDEKLVAK 262
Qy 102 -----KVTHIYLSILGLLILYVLTVEPIIKRLF 134
Db 263 KFKPSYDKLTLMYISLVGFCLSFGWFTLEFPVLFRYFF 302

RESULT 15

S48160

Metalloproteinase (EC 3.4.24.-) H-I precursor - carpet viper

N;Contains: disintegrin

C;Species: *Echis pyramidum leakeyi*

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 21-Jul-2000

C;Accession: S48160
R;Paine, M.J.I.; Moura-Da-Silva, A.M.; Theakston, R.D.G.; Crampton, J.M.
Eur. J. Biochem. 224, 483-488, 1994
A;Title: Cloning of metalloproteinase genes in the carpet viper (*Echis pyramidum leakeyi*)
A;Reference number: S48160; MUID:95010025; PMID:7925363
A;Accession: S48160
A;Molecule type: mRNA
A;Residues: 1-617 <PAI>
A;Cross-references: GB:X78970; NID:g763092; PIDN:CAAS5565.1; PID:g763093
C;Superfamily: mouse meltrin alpha, disintegrin homology
C;Keywords: hydrolase; metalloproteinase; venom
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-192/Domain: propeptide #status predicted <PRO>
F;193-617/Product: metalloproteinase H-I #status predicted <MAT>
F;403-485/Domain: disintegrin homology <DIS>
F;337/Active site: Glu #status predicted
Query Match 7.3%; Score 75.5; DB 2; Length 617;
Best Local Similarity 34.0%; Pred. No.27;
Matches 17; Conservative 6; Mismatches 14; Indels 13; Gaps 2;

Qy 46 CP-----PYKENSCHYKNKISQKDCDCLHVVEPVRG-----PDV 82
Db 482 CELDVFRNGQPYQSNNGYCYNGCPILKNCIHLWKFPAGVNVAPDV 531

Search completed: March 15, 2004, 08:27:47
Job time : 23 secs